

STIC-Biotech/ChemLib

118065

**From:** Chan, Christina  
**Sent:** Monday, March 29, 2004 11:10 AM  
**To:** Yaen, Christopher; STIC-Biotech/ChemLib  
**Subject:** RE: 09883848-RUSH

Please rush. Thanks Chris

*Chris Chan*

TC 1600 New Hire Training Coordinator and SPE 1644  
(571)-272-0841  
Remsen, 3E89

-----Original Message-----

**From:** Yaen, Christopher  
**Sent:** Monday, March 29, 2004 11:01 AM  
**To:** Chan, Christina  
**Subject:** 09883848-RUSH

Hi Christina,

could you please rush the following search:

sequence search on seq id 15, include sequences that are at least 60% identical

thanks

Christopher Yaen  
US Patent Office  
Art Unit 1642  
571-272-0838  
REM 3A20  
REM 3C18

RECEIVED  
MAR 29 2004  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 3/30  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

```
or send an email to license@isb-sib.ch).
-----
EMBL; U26349; AAA95163.1; -
HSSP; Q62226; 1VHH.
MEROPS; C46.UPW; -
InterPro; IPR009045; Hedgehog/DD_pept.
InterPro; IPR003587; Hedgehog_hint N.
InterPro; IPR003586; Hedgehog_hintC.
InterPro; IPR000320; HH signal.
InterPro; IPR006141; InTeinS.
InterPro; IPR001767; Pept_C46_hint.
InterPro; IPR001657; Peptidase_C46.
Pfam; PF01085; HH_signal; 1.
Pfam; PF01079; Hint; 1.
PRINTS; PR00632; SONICHOG.
ProDom; PD003042; HH_signal; 1.
SMART; SM00306; HintN; 1.
SMART; SM00305; HintC; 1.
PROSITE; PS00817; INTENIN_TER; 1.
Signal; Lipoprotein; Palmitate.
Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
Signal; Lipoprotein; Palmitate.
POTENTIAL.
CHAIN 1 22
CHAIN 23 396
CHAIN 23 197
CHAIN 198 396
CHAIN 276 279
DOMAIN 197 198
SITE 267 267
SITE 270 270
ACT_SITE 270 270
LIPID 23 23
LIPID 197 197
SEQUENCE 396 AA; 44087 MW; 7743EC2268A5E9 CRC64;
Query Match 43.7%; Score 1079.5; DB 1; Length 396;
Best Local Similarity 50.1%; Pred. No. 1.2e-70;
Matches 222; Conservative 55; Mismatches 109; Indels 57; Gaps 6;
Qy 6 RCLLLVSVLLVCSGLACGPRG-FGKRHPKLTPLAYKQFPNVAEKTLCASGRYEG 64
Db 5 RIVLAICGGLLVFVRCGGRGPRGVGRRYMRKLVPLHYKQFPNVPKTLGASGSEG 64
Qy 65 KISNSREFKELTNNYPDIIFKDEENTGADRLMTQCKKINALAISVNNQWPGVKLRV 124
Db 65 KIHGSEFIELVNYNPDIIFKDEKTGADRLMTQCKKORVNAISVNNWPGVKLRV 124
Qy 125 TEGWDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH1HCS 184
Db 125 TEGWDEGHHAHDSLHYEGRALDITTSRDRNKYGMRLARLAVEAGFDWVYYESKAH1HVS 184
Qy 185 VKAENSVAKSGCGPPGSAVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDR 244
Db 185 VKADNSLVGRSGGCPFGFAMVMGTGKERKP;SELKIGTVVTTDETQLITSVVLLFLHR 244
Qy 245 DDGAKKVFYIETREPRRLILTAHLLFVAPHNDSATGEPEASGSGPPSGGALGPRAL 304
Db 245 NPYKATFVLEAECHPSKLLVTPNHLFI-----QSSSAGFLP-----T 285
Qy 305 FASRVPCQVVVVAERDGRLLIPAAVHVSVTLSEEAAGAVAPLTAOCTILINVLASY 364
Db 286 FAYRVQIGDLVQIYV---NGTQVSSKVVVRVSL-EEOTGVTPAPMTEHGLLVQGLVSCY 341
Qy 365 AVIEHSHWAHRAFAFRLAHALAALAPARTDRGDSGGGRGGGRVALTAFGAADAP 424
Db 342 ATVESHTLAHVS LAPRLV FQGIASML-----P 368
Qy 425 GAGATAGIHWYSQLLYQIGTWLL 447
Db 369 DLDMSDGVHWYCHILYVLAQYVL 391
```

RESULT 17

DHH2 XENLA

ID DHH2 XENLA STANDARD; PRT; 398 AA.

AC Q91611;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Desert hedgehog protein precursor 2 (DHH-2) (Hedgehog protein 4) (X-HH4).

DE HH4.

GN HH4.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_taxid=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=95401852; PubMed=7671800;

RA Ekker S.C., McGraw L.L., Lai C.-J., Lee J.J., von Kessler D.P., Moon R.T., Beachy P.A.,

RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis."

RL Development 121:2337-2347(1995).

CC -!- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.

CC -!- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).

CC -!- PTM: The C-terminal domain displays an autoprolysis activity and a cholesterol transferase activity. Both activities result in the cleavage of the full-length protein and covalent attachment of a cholesterol moiety to the C-terminal of the newly generated N-terminal fragment (N-product). This covalent modification appears to play an essential role in restricting the spatial distribution of the protein activity to the cell surface. The N-product is the active species in both local and long-range signaling, whereas the C-product has no signaling activity (By similarity).

CC -!- SIMILARITY: belongs to the hedgehog family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; U26350; AAA95164.1; -

HSSP; Q62226; 1VHH.

MEROPS; C46.UPW; -

InterPro; IPR009045; Hedgehog/DD\_pept.

InterPro; IPR003587; Hedgehog\_hint N.

InterPro; IPR003586; Hedgehog\_hintC.

InterPro; IPR000320; HH\_signal.

InterPro; IPR001767; Pept\_C46\_hint.

InterPro; IPR001657; Peptidase\_C46.

Pfam; PF01085; HH\_signal; 1.

Pfam; PF01079; Hint; 1.

PRINTS; PR00632; SONICHOG.

ProDom; PD003042; HH\_signal; 1.

SMART; SM00305; HintN; 1.

SMART; SM00306; HintC; 1.

Developmental protein; Autocatalytic cleavage; Hydrolase; Protease; Signal; Lipoprotein; Palmitate.

POTENTIAL.

SIGNAL 1 23

CHAIN 24 398

CHAIN 24 199

CHAIN 200 398

DOMAIN 278 281

FT DESERT HEDGEHOG PROTEIN 2.

FT DESERT HEDGEHOG PROTEIN 2 N-PRODUCT.

FT DESERT HEDGEHOG PROTEIN 2 C-PRODUCT.

FT POLY-SER.

FT	SITE	199	200	CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	SITE	269	269	INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
FT	ACT_SITE	272	272	ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
FT	LIPID	24	24	N-palmitoyl cysteine (By similarity).
FT	LIPID	199	199	Cholesterol glycine ester (By similarity).
SQ	SEQUENCE	398 AA;	44458 MW;	DBC23AF65F69DD08 CRC64;
Query Match 43.6%; Score 1077.5; DB 1; Length 398;				
Best Local Similarity 48.7%; Pred. No. 1.6e-70;				
Matches 221; Conservative 58; Mismatches 102; Indels 73; Gaps 7;				
QY	1	MLLLARC---	LLLVSSLLVCSGLACPGRG--	FGKRRHPKLTPLAYKQFIPNVAEKT 55
DB	6	ILILAACCCWLLLPV---	-----CCPGRGPGVGGRRYMRLLVPLLYKQFVNPVEKT 57	
QY	56	LGASGRVEGKISRNSERFKELTPNYPDITFKDENTGADRLMTORCKKLNALAI	SVNM 115	
DB	58	LGASGKSEKTRRGSERFKILVPNPNPDITFKDENTGADRLMTERCKORVNALAI	SVNM 117	
QY	116	QWPGVKLVTEGDEDDGHHSESLHYEGRAVDITTSDRDSKYGLMARLAVEAGFDWV	Y 175	
DB	118	MWPGKLVTEGDEDDGHHSDSLHYEGRALDITTSDRDNKYGLMARLAVEAGFDWV	Y 177	
QY	176	ESKAHHCSCVAENSVAAGSGCGPPGSAIVLEGGTKLVKDLSPGDRVLAADOGELLY	235	
DB	178	ESKAHHSVNTDNLGVRSGCGPFGTAMVMETGKKPLSELGLDVTFTTDTGLLIH	237	
QY	236	SDFTFLDRDDGAKVFFVVIETREPRERLILTAHLLFVAFHNDSATGEPEASSGSGPPS	295	
DB	238	SVLFLFHRDYPKATFTVLIIEAGHPYKLVTPNHLFI	-----KS 278	
QY	296	CGALGPRALFASRVPGQ--RVYVAERDGRRLPLPAHVHVSLEBAAGAYAPLTAQGT	353	
DB	279	SSSTGFQPTFAYRVQIGDLIYNGTQVQSKVRVSV-----DEQTGYAPMTSEHT	332	
QY	354	ILINRLVASCVAIVIEHSHWAFAPPLAHALLAALAPARTDRGDSGGDGGCGGRV	413	
DB	333	LLVDGVLTSCATVESHSLTAHSLAPLFLFGIASML-----	369	
QY	414	ALTAFGADAPAGATAGIHWSQLLYQIGTWLL	447	
DB	370	-----PDLHTSDGVHWYCHILYVLAKYVL	393	
RESULT 18				
ID	HH DROME	STANDARD;	PRT;	421 AA.
AC	Q02936; Q9VCQ4;			
DC	01-FEB-1994 (Rel. 28, Created)			
DT	15-MAR-2004 (Rel. 43, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Hedgehog protein precursor.			
GN	HH OR CC4637.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RN	SEQUENCE FROM N.A. (ISOFORM LONG), FUNCTION, DEVELOPMENTAL STAGE, AND			
RP	TISSUE SPECIFICITY.			
RC	STRAIN=Canton-S; TISSUE=Embryo;			
RC	MEDLINE=93185922; PubMed=8166882;			
RA	Takahashi S., Orihara M., Kojima T., Saigo K.;			
RA	"Structure and expression of hedgehog, a Drosophila segment-polarity			
RT	gene required for cell-cell communication."			
RL	Gene 124:183-189(1993).			
RN	[2]			
RN	SEQUENCE FROM N.A. (ISOFORM LONG), FUNCTION, DEVELOPMENTAL STAGE, AND			

RP	TISSUE SPECIFICITY.
RX	MEDLINE=93008241; PubMed=1394430;
RA	Lee J.J., von Kessler D.P., Parks S., Beachy P.A.;
RT	"Secretion and localized transcription suggest a role in positional
RT	signaling for products of the segmentation gene hedgehog."
RL	Cell 71:33-50(1992).
RN	[3]
RN	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, SUBCELLULAR LOCATION,
RP	DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RC	TISSUE=Embryo;
RC	MEDLINE=93083438; PubMed=1280560;
RA	Mohler J., Vani K.;
RT	"Molecular organization and embryonic expression of the hedgehog gene
RT	involved in cell-cell communication in segmental patterning of
RT	Drosophila."
RL	Development 115:957-971(1992).
RN	[4]
RN	SEQUENCE FROM N.A. (ISOFORM LONG), FUNCTION, DEVELOPMENTAL STAGE, AND
RP	TISSUE SPECIFICITY.
RC	STRAIN=Oregon-R; TISSUE=Embryo;
RC	MEDLINE=94040725; PubMed=1340474;
RA	Tabata T., Eaton S., Kornberg T.B.;
RT	"The Drosophila hedgehog gene is expressed specifically in posterior
RT	compartment cells and is a target of engrailed regulation."
RL	Genes Dev. 6:2635-2645(1992).
RN	[5]
RN	SEQUENCE FROM N.A.
RP	STRAIN=Berkley;
RC	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA	Han K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Ballew R.Y., Basu P., Berman B.P., Bhandari D., Bolshakov S.,
RA	Beeson K.M., Bonos P.V., Borker J., Brockstein P., Brotter P.,
RA	Borkova D., Botchan M.R., Bouck J., Butcher H., Cadiou E., Center A.,
RA	Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;"
RT	"The genome sequence of Drosophila melanogaster."
RL	Science 287:2185-2195(2000).
RN	[6]
RN	AUTOPROTEOLYTIC CLEAVAGE.
RP	MEDLINE=95191685; PubMed=7885476;
RA	Porter J.A., von Kessler D.P., Ekker S.C., Young K.E., Lee J.J.,
RA	Moses K., Beachy P.A.;
RT	"The product of hedgehog autoproteolytic cleavage active in local and







DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
DB Sonic hedgehog protein (SHH) (Fragments)  
GN SHH.  
OS Puntius tetrazona (Tiger barb).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Puntius.  
OX NCBI\_TaxID=27709;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=97075114; PubMed=8917540;  
RA Zardoya R., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
closely related to the zebrafish."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE  
NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING  
OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS  
FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)  
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO  
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,  
PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY  
SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
cell, while the N-terminal peptide remains associated with the  
cell surface. Is also secreted in either cleaved or uncleaved form  
to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
and a cholesterol transferase activity. Both activities result in  
the cleavage of the full-length protein and covalent attachment of  
a cholesterol moiety to the C-terminal of the newly generated N-  
terminal fragment (N-product). This covalent modification appears  
to play an essential role in restricting the spatial distribution  
of the protein activity to the cell surface. The N-product is the  
active species in both local and long-range signaling, whereas the  
C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; U51352; AAB38577.1; -  
DR EMBL; U51371; AAB38596.1; -  
DR HSP; Q62226; 1VHH.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR000320; HH signal.  
DR PRINTS; PR00632; SONTCHOG.  
DR ProDom; PD003042; HH signal; 1.  
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
FT NON\_TER 1  
FT NON\_CONS 63 64  
FT NON\_TER 121 121  
SQ SEQUENCE 121 AA; 14003 MW; 559334E4057391C1 CRC64;  
Query Match 24.3%; Score 599; DB 1; Length 121;  
Best Local Similarity 80.9%; Pred No. 1,2e-36;  
Matches 114; Conservative 5; Mismatches 2; Indels 20; Gaps 1;  
QY 30 FGRHRPKKLTPLAYKQFIPNVAEKTLGASGYEKGKISNRSRPFKELPTNPNPDIIFKDE 89  
DB 1 YGRRHRPKKLTPLAYKQFIPNVAEKTLGASGYEKGKISNRSRPFKELPTNPNPDIIFKDE 60  
QY 90 ENTGADRLMTQRCCKLNALATSVNNQWPGVKLRVTEGWDGDEGHHSFSLHYEGRAVDIT 149

Db 61 ENT-----VNNQWPGVKLRVTEGWDGDEGHHSFSLHYEGRAVDIT 100  
QY 150 TSDRDRSKYGMRLARLAVEAGF 170  
DB 101 TSDRDRSKYGMRLARLAVEAGF 121  
RESULT 23  
SHH\_RASHE STANDARD; PRT; 121 AA.  
ID SHH\_RASHE  
AC P79864; P79865;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Sonic hedgehog protein (SHH) (Fragments).  
GN SHH.  
OS Rasbora heteromorpha (Rasbora).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Rasbora.  
OX NCBI\_TaxID=30731;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=97075114; PubMed=8917540;  
RA Zardoya R., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
closely related to the zebrafish."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE  
NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING  
OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS  
FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)  
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO  
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,  
PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY  
SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
cell, while the N-terminal peptide remains associated with the  
cell surface. Is also secreted in either cleaved or uncleaved form  
to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
and a cholesterol transferase activity. Both activities result in  
the cleavage of the full-length protein and covalent attachment of  
a cholesterol moiety to the C-terminal of the newly generated N-  
terminal fragment (N-product). This covalent modification appears  
to play an essential role in restricting the spatial distribution  
of the protein activity to the cell surface. The N-product is the  
active species in both local and long-range signaling, whereas the  
C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; U51354; AAB38579.1; -  
DR EMBL; U51372; AAB38598.1; -  
DR HSP; Q62226; 1VHH.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR000320; HH signal.  
DR InterPro; IPR001657; Peptidase\_C46.  
DR PRINTS; PR00632; SONTCHOG.  
DR ProDom; PD003042; HH signal; 1.  
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
FT NON\_TER 1  
FT NON\_CONS 63 64  
FT NON\_TER 121 121

```
SQ SEQUENCE 121 AA; 13940 MW; 6C9334EA1C73954F CRC64;
Query Match 24.2%; Score 598; DB 1; Length 121;
Best Local Similarity 80.9%; Pred. No. 1.4e-36;
Matches 114; Conservative 5; Mismatches 2; Indels 20; Gaps 1;
QY 30 FGRRHPKLTPLAYKQFIPNVAEKTIGASGRYEGKISRNSERFKELTPNYPNDIIFKDE 89
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 YGRRHPKLTPLAYKQFIPNVAEKTIGASGRYEGKITRNSERFKELTPNYPNDIIFKDE 60
QY 90 ENTGADRLMTORCKDKLNALAISVMNQWPGVKLRVTGWDGHHSESLHYEGRAVDIT 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ENT-----VMMQWPGVKLRVTGWDGHHSESLHYEGRAVDIT 100
QY 150 TSDRDSKYGMRLAVEAGF 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 TSDRDSKYGTLSELAVEAGF 121
RESULT 24
SHH_RASPA STANDARD; PRT; 121 AA.
AC P79869; P79870;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sonic hedgehog protein (SHH) (Fragments).
GN SHH.
OS Rasbora paviei (Sidestripe rasbora).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Rasbora.
OX NCBI_TaxID=38659;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell, while the N-terminal peptide remains associated with the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autotranslation activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
```

```
DR EMBL; U51343; AAB38580.1; -.
DR EMBL; U51362; AAB38599.1; -.
DR HSP; O62226; 1VHH.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT NON_CONS 63 64
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13974 MW; 669334E406739088 CRC64;
Query Match 24.2%; Score 598; DB 1; Length 121;
Best Local Similarity 80.9%; Pred. No. 1.4e-36;
Matches 114; Conservative 5; Mismatches 2; Indels 20; Gaps 1;
QY 30 FGRRHPKLTPLAYKQFIPNVAEKTIGASGRYEGKISRNSERFKELTPNYPNDIIFKDE 89
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 YGRRHPKLTPLAYKQFIPNVAEKTIGASGRYEGKITRNSERFKELTPNYPNDIIFKDE 60
QY 90 ENTGADRLMTORCKDKLNALAISVMNQWPGVKLRVTGWDGHHSESLHYEGRAVDIT 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ENT-----VMMQWPGVKLRVTGWDGHHSESLHYEGRAVDIT 100
QY 150 TSDRDSKYGMRLAVEAGF 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 TSDRDSKYGTLSELAVEAGF 121
RESULT 25
SHH_TANAL STANDARD; PRT; 121 AA.
AC P79915; P79916;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sonic hedgehog protein (SHH) (Fragments).
GN SHH.
OS Tanichthys albonubes (White cloud mountain minnow).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Tanichthys.
OX NCBI_TaxID=38762;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell, while the N-terminal peptide remains associated with the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autotranslation activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
```





QY 150 TSDRDSKYGMARLAVEAGF 170  
DB 101 TSDRDKSKYGLSLAVEAGF 121

RESULT 29  
SHH\_DANAT  
ID SHH\_DANAT STANDARD; PRT; 121 AA.  
AC O13238; O13193; O13209; O13246;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Sonic hedgehog protein (SHH) (Fragments).  
GN SHH.  
OS Danio aff. tweediei.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=46785;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Muscle;  
RZ MEDLINE=97075114; PubMed=8917540;  
RA Zardoya R., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species closely related to the zebrafish."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the cell, while the N-terminal peptide remains associated with the cell surface. Is also secreted in either cleaved or uncleaved form to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity and a cholesterol transferase activity. Both activities result in the cleavage of the full-length protein and covalent attachment of a cholesterol moiety to the C-terminal of the newly generated N-terminal fragment (N-product). This covalent modification appears to play an essential role in restricting the spatial distribution of the protein activity to the cell surface. The N-product is the active species in both local and long-range signaling, whereas the C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U51348; AAB38568.1; -  
CC EMBL; U51367; AAB38585.1; -  
CC HSP; Q62226; LVHH.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR000320; HH\_signal.  
CC InterPro; IPR001657; Peptidase\_C46.  
CC PRINTS; PR00632; SONICHHOG.  
CC ProDom; PD003042; HH\_signal; 1.  
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
FT NON\_TER 1  
FT NON\_TER 63  
FT NON\_TER 121  
FT NON\_TER 121  
SQ SEQUENCE 121 AA; 14012 MW; A58A2DB40573825C CRC64;

Query Match 24.1%; Score 594; DB 1; Length 121;  
Best Local Similarity 80.1%; Pred. No. 2.8e-36;  
Matches 113; Conservative 5; Mismatches 3; Indels 20; Gaps 1;

QY 30 FGRHHPKLTPLAYKQIPNVAETKLGASRGYSGKTSRNSERKELTPNNPDIIKDE 89  
DB 1 YGRRHHPKLTPLAYKQIPNVAETKLGASRGYSGKTSRNSERKELTPNNPDIIKDE 60  
QY 90 ENTGADRLMTQRCOKLNALAISSVNNQWPGVKLVTEGWDGHHSESLHYEGRADIT 149  
DB 61 ENT-----VNNHWPVGKLVTEGWDGHHSESLHYEGRADIT 100  
QY 150 TSDRDSKYGMARLAVEAGF 170  
DB 101 TSDRDKSKYGLSLAVEAGF 121

RESULT 30  
SHH\_DANFR  
ID SHH\_DANFR STANDARD; PRT; 121 AA.  
AC O13245; O13208; O13209; O13246;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Sonic hedgehog protein (SHH) (Fragments).  
GN SHH.  
OS Danio frankei (Leopard danio) (Brachydanio rerio frankei).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=27702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Muscle;  
RZ MEDLINE=97075114; PubMed=8917540;  
RA Zardoya R., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species closely related to the zebrafish."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the cell, while the N-terminal peptide remains associated with the cell surface. Is also secreted in either cleaved or uncleaved form to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity and a cholesterol transferase activity. Both activities result in the cleavage of the full-length protein and covalent attachment of a cholesterol moiety to the C-terminal of the newly generated N-terminal fragment (N-product). This covalent modification appears to play an essential role in restricting the spatial distribution of the protein activity to the cell surface. The N-product is the active species in both local and long-range signaling, whereas the C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U51339; AAB38570.1; -

```
DR EMBL; U51357; AAB38588.1; -.
DR HSSP; Q62226; 1VHH.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 63
FT NON_TER 121
FT NON_TER 121
SQ SEQUENCE 121 AA; 14012 MW; A58A2DE40573B25C CRC64;

Query Match 24.1%; Score 594; DB 1; Length 121;
Best Local Similarity 80.1%; Pred. No. 2.8e-36;
Matches 113; Conservative 5; Mismatches 3; Indels 20; Gaps 1;

QY 30 FGKRRHPKLTPLAYKQFIPNVAEKTGLGASRGYEGKISRSRPFKELTPNYPDIIFKDE 89
Db :|||||
1 YGRRRHPKLTPLAYKQFIPNVAEKTGLGASRGYEGKITNSRPFKELTPNYPDIIFKDE 60

QY 90 ENTGADRLMTQRCCKLNALALSVNMQWPGVKLVTEGWDGHHSESLHYEGRAVDIT 149
Db |||||
61 ENT-----VNNHWPVKLVTEGWDGHHSESLHYEGRAVDIT 100

QY 150 TSDRDRSKYGMRLARLAVEAGF 170
Db |||||
101 TSDRDRSKYGMRLARLAVEAGF 121

RESULT 31
SHH_DANKE STANDARD; PRT; 121 AA.
AC P79709; P79710;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Sonic hedgehog protein (SHH) (Fragments).
GN SHH.
OS Danio rerio (Zebrafish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=38750;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell surface, while the N-terminal peptide remains associated with the
cell surface. It is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -1- PFM: The C-terminal domain displays an autoprotoeolysis activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
```

```
CC C-product has no signaling activity (By similarity).
CC -1- SIMILARITY: Belongs to the hedgehog family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U51340; AAB38571.1; -.
DR EMBL; U51359; AAB38589.1; -.
DR HSSP; Q62226; 1VHH.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 63
FT NON_TER 121
FT NON_TER 121
SQ SEQUENCE 121 AA; 14012 MW; A58A2DE40573B25C CRC64;

Query Match 24.1%; Score 594; DB 1; Length 121;
Best Local Similarity 80.1%; Pred. No. 2.8e-36;
Matches 113; Conservative 5; Mismatches 3; Indels 20; Gaps 1;

QY 30 FGKRRHPKLTPLAYKQFIPNVAEKTGLGASRGYEGKISRSRPFKELTPNYPDIIFKDE 89
Db :|||||
1 YGRRRHPKLTPLAYKQFIPNVAEKTGLGASRGYEGKITNSRPFKELTPNYPDIIFKDE 60

QY 90 ENTGADRLMTQRCCKLNALALSVNMQWPGVKLVTEGWDGHHSESLHYEGRAVDIT 149
Db |||||
61 ENT-----VNNHWPVKLVTEGWDGHHSESLHYEGRAVDIT 100

QY 150 TSDRDRSKYGMRLARLAVEAGF 170
Db |||||
101 TSDRDRSKYGMRLARLAVEAGF 121

RESULT 32
SHH_DANPU STANDARD; PRT; 121 AA.
AC P79717; P79718;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Sonic hedgehog protein (SHH) (Fragments).
GN SHH.
OS Danio pulcher.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=38751;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell surface, while the N-terminal peptide remains associated with the
cell surface. It is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -1- PFM: The C-terminal domain displays an autoprotoeolysis activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
```





DR HSPP: Q62226; 1VHH.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR000320; HH\_signal.  
DR InterPro; IPR001657; Peptidease\_C46.  
DR PRINTS; PRK0832; SONICHHOG.  
DR ProDom; PD003042; HH\_signal; 1.  
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
FT NON\_TER 1  
FT NON\_TER 63  
FT NON\_TER 121  
FT NON\_TER 121  
SQ SEQUENCE 121 AA; 14078 MW; A5832D7405738DAC CRC64;  
  
Query Match 23.9%; Score 590; DB 1; Length 121;  
Best Local Similarity 79.4%; Pred.No.5-4e-36;  
Matches 112; Conservative 6; Mismatches 3; Indels 20; Gaps 1;  
  
QY 30 FGKRHPKKLTPLAYKQFINVAEKTGLASGRYEKGKISNSRFRKELPTNNPDIIFKDE 89  
DB :|::||::|:  
1 YGRRRHPPKKLTPLAYKQFINVAEKTGLASGRYEKGKITNSRFRKELPTNNPDIIFKDE 60  
  
QY 90 ENTGADRLMTQRCKDKLNALAIIVNQWQGVLRTVEGWDEDGHSESSLHYEGRAVDIT 149  
DB 61 ENT-----VNQMFGVKLRMTEGWEDDDHGFESLSHYEGRAVDIT 100  
  
QY 150 TSDRSRKYGMLARLAVERGF 170  
DB ||||::|:  
101 TSDRDKSRYGTLRLAVENGFF 121  
  
RESULT 37  
ID DHH BRARE STANDARD; PRT; 88 AA.  
DT AC F79729; Q9YGUJ;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DN Desert hedgehog protein (DHH) (Fragment).  
GE DHH.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxId=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99051425; PubMed=9831563;  
RA Amores A., Force A., Yan Y.-L.; Wang Y.-L., Fritz A., Prince V.,  
RA Ho R., Amenya C., Langeland J., Westerfield M., Ekker M.,  
RA Postlethwait J.;  
RT "Zebrafish hox clusters and vertebrate genome evolution.";  
RL Science 282:1711-1714(1998).  
RN [2]  
RP SEQUENCE OF 14-71 FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=97075114; PubMed=8917540;  
RA Zardoya K., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
closely related to the zebrafish."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -I- FUNCTION: Intracellular signal essential for a variety of  
cell patterning events during development (By similarity).  
CC -I- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
cell surface. It also secreted in either cleaved or uncleaved form  
to mediate signaling to other calls (By similarity).  
CC -I- PTM: The C-terminal domain displays an autophosphorylation activity  
and a cholesterol transferase activity. Both activities result in  
the cleavage of the full-length protein and covalent attachment of  
cholesterol moiety to the C-terminal of the newly generated N-  
terminal fragment (N-product). This covalent modification appears  
to play an essential role in restricting the spatial distribution  
of the protein activity to the cell surface. The N-product is the  
active species in both local ad long-range signaling, whereas the

DR HSPP: Q62226; 1VHH.

```

CC C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF071236; AAD15931.1; -.
CC EMBL; U51386; AAB38613.1; -.
CC HSSP; Q62226; 1VHH.
CC ZFIN; ZDB-GENE-990714-5; dhh.
CC InterPro; IPR009045; Hedgehog/DD_pept.
CC InterPro; IPR000320; HH_signal.
CC InterPro; IPR001657; Peptidase_C46.
CC Pfam; PF01085; HH_signal; 1.
CC PRINTS; PR00632; SONICHHOG.
CC ProDom; PD003042; HH_signal; 1.
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
CC NON_TER 1
CC FT CONFLICT 17 Q -> M (IN REF. 2).
CC FT CONFLICT 22 R -> K (IN REF. 2).
CC FT CONFLICT 28 A -> G (IN REF. 2).
CC FT CONFLICT 34 38 HPPG -> NLED (IN REF. 2).
CC FT CONFLICT 56 57 TK -> RN (IN REF. 2).
CC FT CONFLICT 61 61 L -> M (IN REF. 2).
CC FT CONFLICT 64 64 Q -> R (IN REF. 2).
CC FT NON_TER 88
CC SQ SEQUENCE 88 AA; 10069 MW; E3D34A0C36677FA6 CRC64;

Query Match 16.0%; Score 396; DB 1; Length 88;
Best Local Similarity 81.8%; Pred. No. 3.6e-22;
Matches 72; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 100 QRCXKLNALAI SVMQWPGVKLVTEGWDEGHSESLHYEGRAVDITTSRDRSKYG 159
DB 1 QRCXKCLYKLAIAVMNQWPGVRLRVTEAWDEGHPPGSLHYEGRAVDITTSRDTKKYG 60

QY 160 MLARLAVEAGFDWVYYSKAHICSVKA 187
DB 61 LLAQLAVEAGFDWVHYESKVVHCSVKA 88

QY 160 MLARLAVEAGFDWVYYSKAHICSVKA 187
DB 61 LLAQLAVEAGFDWVHYESKVVHCSVKA 88

RESULT 38
ID IH CARAU STANDARD; PRT; 58 AA.
AC P79693;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Indian hedgehog protein (IHH) (Fragment).
GN IHH.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OC NCBI_TaxID=7957;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RL closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Intercellular signal essential for a variety of
CC patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autoproteolysis activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of

```

```

CC to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autoproteolysis activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of
CC a cholesterol moiety to the C-terminal of the newly generated N-
CC terminal fragment (N-product). This covalent modification appears
CC to play an essential role in restricting the spatial distribution
CC of the protein activity to the cell surface. The N-product is the
CC active species in both local and long-range signaling, whereas the
CC C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U51386; AAB38601.1; -.
CC HSSP; Q62226; 1VHH.
CC InterPro; IPR009045; Hedgehog/DD_pept.
CC InterPro; IPR000320; HH_signal.
CC InterPro; IPR001657; Peptidase_C46.
CC Pfam; PF01085; HH_signal; 1.
CC PRINTS; PR00632; SONICHHOG.
CC ProDom; PD003042; HH_signal; 1.
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
CC NON_TER 1
CC FT NON_TER 58
CC SQ SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;

Query Match 11.9%; Score 293; DB 1; Length 58;
Best Local Similarity 93.1%; Pred. No. 5.8e-15;
Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 113 VMNQWPGVKLVTEGWDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGF 170
DB 1 VMNLWPGVRLRVTEGWDEGHSESLHYEGRAVDITTSRDRNKYMLARLAVEAGF 58

RESULT 39
ID IH DANAT STANDARD; PRT; 58 AA.
AC O13240; O13200;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Indian hedgehog protein (IHH) (Fragment).
GN IHH.
OS Danio aff. tweediei.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=46785;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RL closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Intercellular signal essential for a variety of
CC patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autoproteolysis activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of

```

CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the  
CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U51381; AAB38602.1; -  
CC HSP; Q62226; 1VHH.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR003020; HH signal.  
CC InterPro; IPR001657; Peptidase\_C46.  
CC Pfam; PF01085; HH signal; 1.  
CC PRINTS; PR00632; SONICHHOG.  
CC ProDom; PD003042; HH signal; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
CC NON\_TER 1  
CC NON\_TER 58  
CC SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;  
CC  
CC Query Match 11.9%; Score 293; DB 1; Length 58;  
CC Best Local Similarity 93.1%; Pred. No. 5.8e-15;  
CC Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
CC  
CC QY 113 VNMQPGVKLVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGF 170  
CC 1 VNMLPGVRLRVTEGWDGHHSESLHYEGRAVDITTSDDRNKYAMRLAVEAGF 58  
CC  
CC RESULT 40  
CC ID IHK\_DANKE STANDARD; PRT; 58 AA.  
CC AC P79711;  
CC DT 15-JUL-1999 (Rel. 38, Created)  
CC DT 15-JUL-1999 (Rel. 38, Last sequence update)  
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
CC DE Indian hedgehog protein (IHH) (Fragment).  
CC GN IHH.  
CC OS Danio rerio (Zebrafish).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
CC OC Cyprinidae; Danio.  
CC OX NCBI\_TaxID=38750;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RC TISSUE=Muscle;  
CC RX MEDLINE=97075114; PubMed=8917540;  
CC RA Zardoya R., Abouheif E., Meyer A.;  
CC RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
CC closely related to the zebrafish."  
CC Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: Intercellular signal essential for a variety of  
CC patterning events during development (By similarity).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
CC cell, while the N-terminal peptide remains associated with the  
CC cell surface. Is also secreted in either cleaved or uncleaved form  
CC to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
CC and a cholesterol transferase activity. Both activities result in  
CC the cleavage of the full-length protein and covalent attachment of  
CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the

CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U51376; AAB38604.1; -  
CC HSP; Q62226; 1VHH.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR003020; HH signal.  
CC InterPro; IPR001657; Peptidase\_C46.  
CC Pfam; PF01085; HH signal; 1.  
CC PRINTS; PR00632; SONICHHOG.  
CC ProDom; PD003042; HH signal; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
CC NON\_TER 1  
CC NON\_TER 58  
CC SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;  
CC  
CC Query Match 11.9%; Score 293; DB 1; Length 58;  
CC Best Local Similarity 93.1%; Pred. No. 5.8e-15;  
CC Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
CC  
CC QY 113 VNMQPGVKLVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGF 170  
CC 1 VNMLPGVRLRVTEGWDGHHSESLHYEGRAVDITTSDDRNKYAMRLAVEAGF 58  
CC  
CC RESULT 41  
CC ID IHK\_DANPU STANDARD; PRT; 58 AA.  
CC AC P79719;  
CC DT 15-JUL-1999 (Rel. 38, Created)  
CC DT 15-JUL-1999 (Rel. 38, Last sequence update)  
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
CC DE Indian hedgehog protein (IHH) (Fragment).  
CC GN IHH.  
CC OS Danio pulcher.  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
CC OC Cyprinidae; Danio.  
CC OX NCBI\_TaxID=38751;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RC TISSUE=Muscle;  
CC RX MEDLINE=97075114; PubMed=8917540;  
CC RA Zardoya R., Abouheif E., Meyer A.;  
CC RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
CC closely related to the zebrafish."  
CC Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: Intercellular signal essential for a variety of  
CC patterning events during development (By similarity).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
CC cell, while the N-terminal peptide remains associated with the  
CC cell surface. Is also secreted in either cleaved or uncleaved form  
CC to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
CC and a cholesterol transferase activity. Both activities result in  
CC the cleavage of the full-length protein and covalent attachment of  
CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the  
CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U51377; AAB38606.1; --  
CC HSPF; Q62226; 1VHH.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR000320; HH\_signal.  
CC Pfam; PF01085; HH\_signal; 1.  
CC PRINTS; PR00632; SONICHOG.  
CC ProDom; PD003042; HH\_signal; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
CC NON\_TER 1  
CC FT NON\_TER 58  
CC SQ SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;  
CC  
CC Query Match 11.9%; Score 293; DB 1; Length 58;  
CC Best Local Similarity 93.1%; Pred. No. 5.8e-15;  
CC Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
CC  
CC QY 113 VMNQWPGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMGLARLAVEAGF 170  
CC DB 1 VMNLWPGVRLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMGLARLAVEAGF 58  
CC  
CC RESULT 42  
CC ID IHH DEVDE STANDARD; PRT; 58 AA.  
CC AC O13243; O13205;  
CC DT 15-JUL-1999 (Rel. 38, Created)  
CC DT 15-JUL-1999 (Rel. 38, Last sequence update)  
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
CC DE Indian hedgehog protein (IH) (Fragment).  
CC GN IH.  
CC OS Devario devario (Danio devario).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
CC OC Cyprinidae; Danio.  
CC OX NCBI\_TaxID=46781;  
CC RN [1]\_TaxID=46781;  
CC RP SEQUENCE FROM N.A.  
CC RC TISSUE=Muscle;  
CC RX MEDLINE=97075114; PubMed=8917540;  
CC RA Zardoya R., Abouheif E., Meyer A.;  
CC RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
CC closely related to the zebrafish."  
CC Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC RL  
CC -!- FUNCTION: Intercellular signal essential for a variety of  
CC patterning events during development (By similarity).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
CC cell, while the N-terminal peptide remains associated with the  
CC cell surface. Is also secreted in either cleaved or uncleaved form  
CC to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
CC and a cholesterol transferase activity. Both activities result in  
CC the cleavage of the full-length protein and covalent attachment of  
CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the  
CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U51382; AAB38603.1; --  
CC HSPF; Q62226; 1VHH.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR000320; HH\_signal.  
CC Pfam; PF01085; HH\_signal; 1.  
CC PRINTS; PR00632; SONICHOG.  
CC ProDom; PD003042; HH\_signal; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
CC NON\_TER 1  
CC FT NON\_TER 58  
CC SQ SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;  
CC  
CC Query Match 11.9%; Score 293; DB 1; Length 58;  
CC Best Local Similarity 93.1%; Pred. No. 5.8e-15;  
CC Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
CC  
CC QY 113 VMNQWPGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMGLARLAVEAGF 170  
CC DB 1 VMNLWPGVRLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMGLARLAVEAGF 58  
CC  
CC RESULT 43  
CC ID IHH PUNTE STANDARD; PRT; 58 AA.  
CC AC P79852;  
CC DT 15-JUL-1999 (Rel. 38, Created)  
CC DT 15-JUL-1999 (Rel. 38, Last sequence update)  
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
CC DE Indian hedgehog protein (IH) (Fragment).  
CC GN IH.  
CC OS Puntius tetrazona (Tiger barb).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
CC OC Cyprinidae; Puntius.  
CC OX NCBI\_TaxID=27709;  
CC RN [1]\_TaxID=27709;  
CC RP SEQUENCE FROM N.A.  
CC RC TISSUE=Muscle;  
CC RX MEDLINE=97075114; PubMed=8917540;  
CC RA Zardoya R., Abouheif E., Meyer A.;  
CC RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
CC closely related to the zebrafish."  
CC Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC RL  
CC -!- FUNCTION: Intercellular signal essential for a variety of  
CC patterning events during development (By similarity).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
CC cell, while the N-terminal peptide remains associated with the  
CC cell surface. Is also secreted in either cleaved or uncleaved form  
CC to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
CC and a cholesterol transferase activity. Both activities result in  
CC the cleavage of the full-length protein and covalent attachment of  
CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the  
CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
DR EMBL; U51384; AAB38608.1; -.
DR HSP; Q62226; 1VHH.
DR InterPro; IPR000045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
DR ProDom; PD003042; HH_signal; 1.
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 58
SQ SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;
Query Match 11.9%; Score 293; DB 1; Length 58;
Best Local Similarity 93.1%; Pred. No. 5.8e-15;
Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 113 VMNQPGVKLVTEGWEDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGF 170
Db 1 VMNLMPGVLRLVTEGWEDGHHSSESLHYEGRAVDITTSDDRNKYAMLARLAVEAGF 58
RESULT 44
IHH_RASEL STANDARD; PRT; 58 AA.
AC P79860;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Indian hedgehog protein (IHH) (Fragment).
GN IHH.
OS Rasbora elegans (Elegant rasbora).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Rasbora.
OX NCBI_TaxID=27712;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RT closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Interleukin signal essential for a variety of
CC patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
CC cell, while the N-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autoprolysis activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of
CC a cholesterol moiety to the C-terminal of the newly generated N-
CC terminal fragment (N-product). This covalent modification appears
CC to play an essential role in restricting the spatial distribution
CC of the protein activity to the cell surface. The N-product is the
CC active species in both local and long-range signaling, whereas the
CC C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51385; AAB38609.1; -.
CC HSP; Q62226; 1VHH.
DR InterPro; IPR000045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
```

```
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 58
SQ SEQUENCE 58 AA; 6658 MW; 2CC8F53CEC04D809 CRC64;
Query Match 11.9%; Score 293; DB 1; Length 58;
Best Local Similarity 93.1%; Pred. No. 5.8e-15;
Matches 54; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 113 VMNQPGVKLVTEGWEDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGF 170
Db 1 VMNLMPGVLRLVTEGWEDGHHSSESLHYEGRAVDITTSDDRNKYAMLARLAVEAGF 58
RESULT 45
SHH_PSEPR STANDARD; PRT; 58 AA.
AC P79839;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Sonic hedgehog protein (SHH) (Fragment).
GN SHH.
OS Pseudorasbora parva.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Pseudorasbora.
OX NCBI_TaxID=51549;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RT closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC NOTOCHORD EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC OF THE BRAIN AND EARLY PATTERNING OF THE DORSO-VENTRAL PATTERNING
CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
CC PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
CC cell, while the N-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autoprolysis activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of
CC a cholesterol moiety to the C-terminal of the newly generated N-
CC terminal fragment (N-product). This covalent modification appears
CC to play an essential role in restricting the spatial distribution
CC of the protein activity to the cell surface. The N-product is the
CC active species in both local and long-range signaling, whereas the
CC C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51358; AAB38594.1; -.
CC
CC DR
```

```
DR HSSP: Q62226; 1VHH.
DR InterPro: IPR009045; Hedgehog/DD_pept.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; Peptidase_C46.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHOG.
DR ProDom: PD003042; HH_signal; 1.
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT SEQUENCE 58 AA; 6608 MW; BDEAB71BD19FA6B9 CRC64;
Query Match 11.7%; Score 289; DB 1; Length 58;
Best Local Similarity 91.4%; Pred. No. 1.1e-14;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 113 VMNQPGVKLVRTGWDGHHSESLHYEGRVADITTSDDRSKYGMRLAIVEAGF 170
DB 1 VMNQPGVKLVRTGWDGHHSESLHYEGRVADITTSDDRSKYGMRLAIVEAGF 58

RESULT 46
DHH_DANKE
ID DHH_DANKE STANDARD; PRT; 58 AA.
AC P79712;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Desert hedgehog protein (DHH) (Fragment).
GN DHH.
OS Danio kerri (Kerri's danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=38750;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Inter-cellular signal essential for a variety of
patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHOG.
DR ProDom: PD003042; HH_signal; 1.
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON_TER 1
FT SEQUENCE 58 AA; 6657 MW; 62DBC28ABC2A5390 CRC64;
Query Match 11.7%; Score 288; DB 1; Length 58;
Best Local Similarity 91.4%; Pred. No. 1.3e-14;
Matches 53; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 113 VMNQPGVKLVRTGWDGHHSESLHYEGRVADITTSDDRSKYGMRLAIVEAGF 170
DB 1 VMNQPGVKLVRTGWDGHHSESLHYEGRVADITTSDDRSKYGMRLAIVEAGF 58

RESULT 47
IHH_DEVMA
ID IHH_DEVMA STANDARD; PRT; 58 AA.
AC O13215;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Indian hedgehog protein (IHH) (Fragment).
GN IHH.
OS Devario malabaricus (Danio malabaricus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=46780;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Inter-cellular signal essential for a variety of
patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```

FT NON TER 1 1
FT NON TER 58
SQ SEQUENCE 58 AA; 6714 MW; 2CC8E2EED9D50809 CRC64;

Query Match
Best Local Similarity 11.7%; Score 288; DB 1; Length 58;
Matches 52; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 113 VNMWPGVKLVRTGWDGDGHHSSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 170
Db 1 VNMWPGVKLVRTGWDGDGHHSSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 58

RESULT 48
ID IHH_RASHE STANDARD; PRT; 58 AA.
AC P79866;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Indian hedgehog protein (IHH) (Fragment).
GN IHH.
OS Rarbora heteromorphs (Rarbora).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Rasbora.
OX NCBI_TaxID=30731;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Inter cellular signal essential for a variety of
patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell, while the N-terminal peptide remains associated with the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U51383; AAB38610.1; -.
CC HSP; Q62226; 1VHH.
CC InterPro; IPR009045; Hedgehog/DD_pept.
CC InterPro; IPR000320; HH_signal.
CC InterPro; IPR001657; Peptidase_C46.
CC Pfam; PF01085; HH_signal; 1.
CC PRINTS; PR00632; SONICHHOG.
CC ProDom; PD003042; HH_signal; 1.
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON TER 1 1
FT NON TER 58
SQ SEQUENCE 58 AA; 6628 MW; 2CC8F53CF7056809 CRC64;

```

```

Query Match
Best Local Similarity 11.7%; Score 288; DB 1; Length 58;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 113 VNMWPGVKLVRTGWDGDGHHSSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 170
Db 1 VNMWPGVKLVRTGWDGDGHHSSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 58

RESULT 49
ID IHH_DEVPA STANDARD; PRT; 58 AA.
AC O13220;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Indian hedgehog protein (IHH) (Fragment).
GN IHH.
OS Devario pathirana (Barred danio) (Danio pathirana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=46779;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -!- FUNCTION: Inter cellular signal essential for a variety of
patterning events during development (By similarity).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell, while the N-terminal peptide remains associated with the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
CC -!- PTM: The C-terminal domain displays an autophosphorylation activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U51379; AAB38607.1; -.
CC HSP; Q62226; 1VHH.
CC InterPro; IPR009045; Hedgehog/DD_pept.
CC InterPro; IPR000320; HH_signal.
CC InterPro; IPR001657; Peptidase_C46.
CC Pfam; PF01085; HH_signal; 1.
CC PRINTS; PR00632; SONICHHOG.
CC ProDom; PD003042; HH_signal; 1.
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.
FT NON TER 1 1
FT NON TER 58
SQ SEQUENCE 58 AA; 6672 MW; 2CCD5353C04D809 CRC64;

```

```

Query Match
Best Local Similarity 11.6%; Score 287; DB 1; Length 58;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Search completed: March 29, 2004, 18:32:32  
Job time : 21 secs

Qy 113 VMNWPQVGLRVTEGWDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 170  
Db 1 VMNLPQVGLRVTEGWDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 58

RESULT 50  
ID\_IHH\_RASPA STANDARD; PRT; 58 AA.  
AC P79871;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Indian hedgehog protein (IHH) (Fragment).  
GN IHH.  
OS Rasbora paviei (Sidestripe rasbora).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Rasbora.  
OX NCBI\_TaxID=38659;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RA MEDLINE=97075114; PubMed=8917540;  
RX Zardoya R., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
closely related to the zebrafish.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: Intercellular signal essential for a variety of  
patterning events during development (By similarity).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
cell, while the N-terminal peptide remains associated with the  
cell surface. It is also secreted in either cleaved or uncleaved form  
to mediate signaling to other cells (By similarity).  
CC -!- PTM: The C-terminal domain displays an autotransmembrane activity  
and a cholesterol transferase activity. Both activities result in  
the cleavage of the full-length protein and covalent attachment of  
a cholesterol moiety to the C-terminal of the newly generated N-  
terminal fragment (N-product). This covalent modification appears  
to play an essential role in restricting the spatial distribution  
of the protein activity to the cell surface. The N-product is the  
active species in both local and long-range signaling, whereas the  
C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U51378; AB38611.1; --  
CC HSPF; Q62226; 1VHH.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR000320; HH\_signal.  
CC InterPro; IPR001657; Peptidase\_C46.  
CC Pfam; PF01085; HH\_signal; 1.  
CC PRINTS; PR00632; SONICHHOG.  
CC ProDom; PD003042; HH\_signal; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease.  
FT NON\_TER 1  
FT NON\_TER 58  
SQ SEQUENCE 58 AA; 6688 MW; A43BF53CEC1A1735 CRC64;

Query Match 11.6%; Score 287; DB 1; Length 58;  
Best Local Similarity 91.4%; Pred. No. 1.6e-14;  
Matches 53; Conservative 2; Mismatches 3; Indels 0; Caps 0;  
Qy 113 VMNWPQVGLRVTEGWDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 170  
Db 1 VMNLPQVGLRVTEGWDEGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGF 58

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 18:25:24 ; Search time 45 Seconds

(without alignments)

3330.470 Million cell updates/sec

Title: US-09-883-848A-15

Perfect score: 2469

Sequence: 1 MLLARCLLVVSSLLVCS.....GMVKSXSRGAGGAREGA 475

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

- 1: SPREMBL.25.\*
- 2: sp\_archaea.\*
- 3: sp\_bacteria.\*
- 4: sp\_fungi.\*
- 5: sp\_human.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2062	83.5	437	11	Q8C765
2	1544.5	62.6	434	13	O57404
3	1466.5	59.4	413	13	O8AXT0
4	1454	58.9	414	13	Q9W7Q9
5	1308	53.0	406	13	O57567
6	1262	51.1	449	11	Q8OXI9
7	1094.5	44.3	415	5	Q9U526
8	1084.5	43.9	415	5	O17499
9	1049.5	42.5	442	13	O73803
10	1041.5	42.2	447	5	Q9G8A8
11	1010	40.9	381	5	Q7Z264
12	995	40.3	421	5	Q7Z263
13	938	38.0	416	5	Q8MWG6
14	916	37.1	410	5	O61676
15	850	34.4	161	11	Q9R179
16	821	33.3	402	5	Q8MY56

17	758	30.7	557	5	Q8MY57
18	758	30.7	1073	5	Q8MUM7
19	733	29.7	150	13	Q9YGV7
20	731	29.6	139	6	Q9XSI6
21	709	28.7	138	13	Q9W6C1
22	702	28.4	177	11	Q9WV29
23	649	26.3	121	6	Q8MUV7
24	619	25.1	185	5	Q96699
25	531	21.5	99	13	Q8QGN4
26	526.5	21.3	202	6	Q8MIN9
27	523	21.2	119	13	Q42128
28	511	20.7	129	11	Q9WUP6
29	476.5	19.3	137	13	Q42234
30	392	15.9	80	13	Q42441
31	392	15.9	87	5	Q9TX30
32	339	13.7	63	13	Q9R184
33	332	13.4	64	13	Q9PRF5
34	298	12.1	60	5	Q952D8
35	289	11.7	56	13	Q9O2R0
36	245	9.9	74	6	Q866A4
37	237	9.6	49	5	Q9TX33
38	229	9.3	49	5	Q9TX31
39	188	7.6	54	13	Q42233
40	187.5	7.6	125	11	Q9ESH3
41	175.5	7.1	48	5	Q9TX32
42	159.5	6.5	615	5	P91573
43	149	6.0	1207	5	Q21535
44	138.5	5.6	790	5	Q22872
45	138.5	5.6	1137	5	Q21835
46	138.5	5.6	1160	5	Q7YXC8
47	134.5	5.4	481	5	O45992
48	136.5	5.1	1021	5	Q9XUV2
49	133.5	5.0	2639	5	O76786
50	122	4.9	702	6	Q863A2
51	121	4.9	946	5	Q9XV14
52	120.5	4.9	702	4	Q7Z647
53	119	4.8	318	16	O86703
54	119	4.8	2655	5	Q964F4
55	118	4.8	1832	3	O8TGH8
56	116	4.7	1329	16	O06810
57	116	4.7	1408	16	Q8VK17
58	115	4.7	486	16	Q9AAT9
59	115	4.7	795	16	Q7U020
60	115	4.7	1408	16	Q7U022
61	114.5	4.6	652	2	Q44100
62	114	4.6	738	16	Q8VK15
63	114	4.6	741	16	O06808
64	113.5	4.6	753	16	Q811P3
65	113.5	4.6	991	2	Q848A4
66	113.5	4.6	1318	5	Q9NE11
67	113.5	4.6	2172	2	Q7X4R4
68	113	4.6	803	10	Q7XJJ6
69	113	4.6	3324	2	Q8KUP6
70	112.5	4.6	623	2	O9APG6
71	112.5	4.6	1463	16	Q9ADP6
72	112	4.5	513	6	Q7VQC9
73	112	4.5	893	5	Q9W443
74	112	4.5	912	5	Q9BIT2
75	112	4.5	944	5	Q8IRR2
76	111.5	4.5	209	2	Q84912
77	111.5	4.5	1016	5	Q9U0V3
78	111	4.5	632	16	Q8CJT6
79	111	4.5	961	16	Q9EWD3
80	110.5	4.5	471	16	Q9EWD3
81	110.5	4.5	624	16	Q8XZA9
82	110.5	4.5	729	16	Q8P3M3
83	110.5	4.5	1294	16	Q8PKH5
84	110.5	4.5	11096	2	Q9L4W3
85	110	4.5	429	3	Q872V8
86	109.5	4.4	1726	16	Q7WKT6
87	109.5	4.4	1729	16	Q7WJE7
88	109	4.4	398	16	Q7WEG4
89	109	4.4	436	5	Q967T8

Q8MY57	ciona intes
Q8MUM7	helobdella
Q9YGV7	ambystoma m
Q9XSI6	bos taurus
Q9W6C1	eleutheroda
Q9WV29	rattus norv
Q8MUV7	juncus muri
Q96699	anomia coe
Q8QGN4	scyllorhinu
Q8MIN9	equus cabal
Q42128	oryzias lat
Q9WUP6	rattus norv
Q42234	coturnix co
Q42441	oryzias lat
Q9TX30	anopheles g
Q9R184	rana carebb
Q9PRF5	oryzias lat
Q952D8	hydra atten
Q9O2R0	oryzias lat
Q866A4	sus scrofa
Q9TX33	hirudo medi
Q9TX31	tribolium c
Q42233	coturnix co
Q9ESH3	rattus norv
Q9TX32	strongyloce
P91573	caenorhabdi
Q21535	caenorhabdi
Q22872	caenorhabdi
Q21835	caenorhabdi
Q7YXC8	caenorhabdi
O45992	caenorhabdi
Q9XUV2	caenorhabdi
O76786	antheraea p
Q863A2	canis famil
Q9XV14	caenorhabdi
Q7Z647	homo sapien
O86703	streptomyce
Q964F4	antheraea y
Q8TGH8	podospora a
O06810	mycobacteri
Q8VK17	mycobacteri
Q9AAT9	caulobacter
Q7U020	mycobacteri
Q7U022	mycobacteri
Q44100	anaplasma m
O8VK15	mycobacteri
O06808	mycobacteri
Q811P3	bacillus ce
Q848A4	streptomyce
Q9NE11	leishmania
Q7X4R4	streptomyce
Q8KUP6	actinosyne
Q9APG6	anaplasma m
Q9ADP6	streptomyce
Q7VQC9	canis famil
Q9W443	drosophila
Q9BIT2	plectreury
Q8IRR2	drosophila
Q84912	streptomyce
Q9U0V3	leishmania
Q8CJT6	streptomyce
Q9EWD3	streptomyce
Q9EWD3	streptomyce
Q8XZA9	raistonia s
Q8P3M3	xanthomonas
Q8PKH5	xanthomonas
Q9L4W3	streptomyce
Q872V8	neurospora
Q7WKT6	bordetella
Q7WJE7	bordetella
Q7WEG4	bordetella
Q967T8	antheraea p

```

90 109 4.4 482 4 Q8TAT1
91 109 4.4 587 16 Q8SRL0
92 109 4.4 605 5 Q8MRE7
93 108.5 4.4 202 10 Q8S7J6
94 108.5 4.4 202 10 Q7XF88
95 108.5 4.4 795 10 Q9LJ20
96 108 4.4 803 16 Q81ZM2
97 108 4.4 850 16 Q9LOV2
98 108 4.4 1048 2 Q84CG4
99 108 4.4 1531 2 Q93N88
100 107.5 4.4 5192 2 Q93TW9
101 107 4.3 313 16 Q8FUB7
102 106.5 4.3 602 16 Q8XBT5
103 106.5 4.3 655 16 Q9FBR7
104 106.5 4.3 2675 2 Q9FB23
105 106 4.3 327 10 Q9ART6
106 106 4.3 673 16 Q82H40
107 106 4.3 1067 5 Q7YU69
108 106 4.3 2902 2 Q7WTF5
109 106 4.3 10917 2 Q93NM6
110 105.5 4.3 490 10 Q04270
111 105.5 4.3 493 5 Q8T7X6
112 105.5 4.3 698 12 Q806B2
113 105.5 4.3 841 16 Q9AAE0
114 105.5 4.3 893 16 Q81ZX0
115 105.5 4.3 1324 2 Q52048
116 105.5 4.3 7463 16 Q924X6
117 105 4.3 387 16 Q82C63
118 105 4.3 398 16 Q7W347
119 105 4.3 437 10 Q7Y1N8
120 105 4.3 460 5 Q9GF2F
121 105 4.3 890 2 Q53592
122 105 4.3 1729 16 Q7VVC2
123 105 4.3 3738 2 Q8G8A6
124 104.5 4.2 464 16 Q82L07
125 104 4.2 240 16 Q89XP1
126 104 4.2 304 10 Q84R72
127 104 4.2 519 5 Q18350
128 104 4.2 519 5 Q18536
129 104 4.2 821 2 Q8VPM9
130 104 4.2 3161 2 Q54511
131 104 4.2 3579 2 Q8GBX4
132 103.5 4.2 481 16 Q8VXM5
133 103.5 4.2 1952 2 Q9L4W4
134 103 4.2 195 5 Q23193
135 103 4.2 457 16 Q82IK3
136 103 4.2 664 5 Q9NEC7
137 103 4.2 717 16 Q9L0I2
138 103 4.2 1907 16 Q93HI9
139 102.5 4.2 413 16 P95236
140 102.5 4.2 413 16 Q7TYS4
141 102.5 4.2 479 16 Q06292
142 102.5 4.2 479 16 Q7U280
143 102.5 4.2 742 16 Q812X5
144 102.5 4.2 769 10 Q9LLJ1
145 102.5 4.2 1346 10 Q8L640
146 102.5 4.2 1742 16 Q8P377
147 102.5 4.2 2090 5 Q9W2T1
148 102.5 4.2 3979 2 Q7WTF1
149 102 4.1 607 4 Q8WVVO
150 102 4.1 623 2 Q9APG7

```

## ALIGNMENTS

```

RESULT 1
Q8C765 PRELIMINARY; PRT; 437 AA.
ID Q8C765
AC Q8C765;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

```

```

DE Sonic hedgehog homolog.
GN SHH OR 9530036O11RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RC MEDLINE=22354683; PubMed=12468851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK052447; BAC34996.1; -.
DR PIR; PT0645; PT0645.
DR MGD; MGI:98297; Shh.
DR GO; GO:0007411; P:axon guidance; IDA.
DR GO; GO:0001708; P:cell fate specification; IMP.
DR GO; GO:0007417; P:central nervous system development; IMP.
DR GO; GO:0007417; P:eye morphogenesis (sensu Mammalia); IMP.
DR GO; GO:0001654; P:eye morphogenesis; IDA.
DR GO; GO:0030539; P:male genital morphogenesis; IMP.
DR GO; GO:0007389; P:pattern specification; IMP.
DR GO; GO:0042127; P:regulation of cell proliferation; IDA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR003586; Hedgehog_hintC.
DR InterPro; IPR003587; Hedgehog_hintN.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR006141; Intein_S.
DR InterPro; IPR001657; Peptidase_C46.
DR InterPro; IPR01767; Pept_C46_hint.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR PROSITE; PS50817; INTEIN_N_TER; 1.
SQ SEQUENCE 437 AA; 47788 MW; 49F1503875BB4227 CRC64;

Query Match 83.5%; Score 2062; DB 11; Length 437;
Best Local Similarity 87.2%; Pred. No. 3.2e-145;
Matches 403; Conservative 8; Mismatches 25; Indels 26; Gaps 3;

QY 1 MLLARCLLLVLVSSLLVCSGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGASG 60
Db L L L L L A R C F L V I L A S S L L V C P G L A C G P G R F G K R R H P K K L T P L A Y K Q F I P N V A E K T I G A S G 61
QY 61 RYEGKISRNSRPFKELTNNYNDPIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120
Db R Y E G K I T R N S R F K E L T N Y N P D I I F K D E E N T G A D R L M T Q C K D K L N A L A I S V M N Q P G V 121
QY 121 KLRVTEGDEDDHSEESLHYEGRAVDITTSDRRSKYGMLARLAVEAGFDWVYYESKAH 180
Db K L R V T E G D E D D H S E E S L H Y E G R A V D I T T S D R R S K Y G M L A R L A V E A G F D W V Y Y E S K A H 181
QY 181 IHCSVKAENSVAAGSGGCPGPGSATVHLQGGTKLVKOLSPGDRVLAADQGRLLYSDFLT 240
Db I H C S V K A E N S V A A G S G G C P G P G S A T V H L Q G G T K L V K O L S P G D R V L A A D Q G R L L Y S D F L T 241
QY 241 FLDRODGAKKVYVETETREPRERLLTTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
Db F L D R O D G A K K V Y V E T E T R E P R E R L L T T A A H L L F V A P H N D S A T G E P E A S S G S G P P S G G A L G 300
QY 301 PRALFASVRPQQRVYVVAERDGRRLLPAAVHVSVTLSSEAGAYAPLTAOITLINRVL 360
Db P R A L F A S V R P Q Q R V Y V V A E R D G R R L L P A A V H V S V T L S E E A G A Y A P L T A O I T L I N R V L 360
QY 361 ASCYAVIEEHSWAHFAFAFRLAHALLAALAPARTDRGDSGGDRGGGRGGRVALTAPCA 420
Db A S C Y A V I E E H S W A H F A F A F R L A H A L L A A L A P A R T D R G D S G G D R G G R G G R V A L T A P C A 420
QY 420 PSALFASVRPQQRVYVVAERDGRRLLPAAVHVSVTLSSEAGAYAPLTAHGTILINRVL 346
Db P S A L F A S V R P Q Q R V Y V V A E R D G R R L L P A A V H V S V T L S E E B A G A Y A P L T A H G T I L I N R V L 346
QY 437 ASCYAVIEEHSWAHFAFAFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
Db A S C Y A V I E E H S W A H F A F A F R L A H A L L A A L A P A R T D ----- G G G G G S I P - A A Q S A 395

```







01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
Hedgehog protein.  
GN HEDGEHOG.  
OS Branchiostoma belcheri (Amphioxus).  
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
OC Branchiostoma.  
NCBI\_TaxID=7741;  
RN [1]  
PP SEQUENCE FROM N.A.  
RA Mao B., Zhang Y., Zhang S., Zhang H.;  
RT "A preliminary study on Qingdao amphioxus hedgehog gene.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ245882; CAB58231.2; -.  
DR HSSP; Q62226; 1VHH.  
DR GO; GO:0008233; P:peptidase activity; IEA.  
DR GO; GO:0007267; P:cell-cell signaling; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0006508; P:protein splicing; IEA.  
DR GO; GO:0006508; P:protein splicing; IEA.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR003586; Hedgehog hintC.  
DR InterPro; IPR003587; Hedgehog hint\_N.  
DR InterPro; IPR00320; HH signal.  
DR InterPro; IPR006141; Inein S.  
DR InterPro; IPR001657; Peptidase C46.  
DR InterPro; IPR001767; Pept\_C46\_hint.  
DR Pfam; PF01085; HH\_signal; 1.  
DR Pfam; PF01079; Hint; 1.  
DR PRINTS; PR00632; SONICHHOG.  
DR ProDom; PD003042; HH\_signal; 1.  
DR SMART; SM00305; HintC; 1.  
DR SMART; SM00306; HintN; 1.  
DR PROSITE; PS50817; INTEIN\_N\_TER; 1.  
SQ SEQUENCE 415 AA; 46912 MW; BF864704752BD6A8 CRC64;

Query Match 44.3%; Score 1094.5; DB 5; Length 415;  
Best Local Similarity 50.2%; Pred. No. 3.3e-73;  
Matches 231; Conservative 57; Mismatches 111; Indels 61; Gaps 10;  
QY 3 LIARCLL-LVIVSSLLVCSG--LACGPGRGKGRHPKLTPLAYKQIPNVAEKLIGAS 59  
DB 4 VLARWVTLVAITLGTHTMGSEACGPGTGRRRHRPKLTPFVYKQMPAVSENTFGAS 63  
QY 60 GRYEGKISNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQWPG 119  
DB 64 GLFNGRITDSRFFHTLKQNFNTDIIFKDEKTDGDRFMTORCKDKLNALAI SVMNQWEG 123  
QY 120 VKLRVTEGWDEGDHSHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKA 179  
DB 124 VKLRVTEGWDEGDHSHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKA 183  
QY 180 HHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDQGRLLYSDFL 239  
DB 184 HHCSVKAESDATTATQGGCFPEESWTRDDGKRIRMRDVRPGDKVLSNDGAGHVFSEVL 243  
QY 240 TFLDRDDGAKKYFYVITETREPRERLLLTAHLLFVAPHNDSATGPEASGSGPPSGGAL 299  
DB 244 TFMDRDSRGFWVYITIKTDRNITVTATPSHLVFTESRDLS----- 285  
QY 300 GPR-ALFASVRPGQVYVVAERDGD--RRLPAAVHVSVTLSEBAAGAYAPLTAQGTILNI 356  
DB 286 APRVAKFMSDVRPGE-FILTPESDGGGFEKV---EIVSVTWREK-KGAYAPLTHGIVVV 340  
QY 357 NRVLASCYAVITEHSAHRAFAFPLRLAHALLAALAPARTDRGSGSGGDRGGGGRVALT 416  
DB 341 DNVMSCYALIESQALAHVWFAFRLYYQLTSSL----- 374  
QY 417 APGADACAGAT--AGTHWYSQLYQIGTWLDSALHP 454  
DB 375 ---WDGSHDQTLQEGVHWYPSFFRYGISLVEPTLLHP 410

RESULT 8  
017499 PRELIMINARY; PRT; 415 AA.  
ID O17499  
AC O17499;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Amphihh protein.  
GN AMPHIHH.  
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).  
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
OC Branchiostoma.  
NCBI\_TaxID=7739;  
RN [1]  
PP SEQUENCE FROM N.A.  
RA Shimeld S.M.;  
RT "Sequence and expression of an amphioxus hedgehog gene.";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y13858; CAA74169.1; -.  
DR HSSP; Q62226; 1VHH.  
DR GO; GO:0008233; P:peptidase activity; IEA.  
DR GO; GO:0007267; P:cell-cell signaling; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0006508; P:protein splicing; IEA.  
DR GO; GO:0006508; P:protein splicing; IEA.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR003586; Hedgehog hintC.  
DR InterPro; IPR003587; Hedgehog hint\_N.  
DR InterPro; IPR00320; HH\_signal.  
DR InterPro; IPR006141; Inein S.  
DR InterPro; IPR001657; Peptidase C46.  
DR InterPro; IPR001767; Pept\_C46\_hint.  
DR Pfam; PF01085; HH\_signal; 1.  
DR Pfam; PF01079; Hint; 1.  
DR PRINTS; PR00632; SONICHHOG.  
DR ProDom; PD003042; HH\_signal; 1.  
DR SMART; SM00305; HintC; 1.  
DR SMART; SM00306; HintN; 1.  
DR PROSITE; PS50817; INTEIN\_N\_TER; 1.  
SQ SEQUENCE 415 AA; 46765 MW; 5E8D2BD2158DD931 CRC64;  
Query Match 43.9%; Score 1084.5; DB 5; Length 415;  
Best Local Similarity 50.0%; Pred. No. 1.8e-72;  
Matches 229; Conservative 53; Mismatches 119; Indels 57; Gaps 8;  
QY 3 LIARCLL-LVIVSSLLVCSG--LACGPGRGKGRHPKLTPLAYKQIPNVAEKLIGAS 59  
DB 4 VLARWVTLVAITLGTHTMGSEACGPGTGRRRHRPKLTPFVYKQMPAVSENTFGAS 63  
QY 60 GRYEGKISNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQWPG 119  
DB 64 GLFNGRITDSRFFHTLKQNFNTDIIFKDEKTDGDRFMTORCKDKLNALAI SVMNQWEG 123  
QY 120 VKLRVTEGWDEGDHSHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKA 179  
DB 124 VKLRVTEGWDEGDHSHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKA 183  
QY 180 HHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDQGRLLYSDFL 239  
DB 184 HHCSVKAESDATTATQGGCFPEESWTRDDGKRIRMRDVRPGDKVLSNDGAGHVFSEVL 243  
QY 240 TFLDRDDGAKKYFYVITETREPRERLLLTAHLLFVAPHNDSATGPEASGSGPPSGGAL 299  
DB 244 TFMDRDSRGFWVYITIKTDRNITVTATPSHLVFTESRDLS----- 286  
QY 300 GPR-ALFASVRPGQVYVVAERDGD--RRLPAAVHVSVTLSEBAAGAYAPLTAQGTILNI 358  
DB 287 -PRIAKFMSDVRPGE-FILTPESDGGGFEKV---EIVSVTWREK-KGAYAPLTHGIVVDN 342  
QY 359 VLASCYAVITEHSAHRAFAFPLRLAHALLAALAPARTDRGSGSGGDRGGGGRVALTAP 418  
DB 343 VAMSCYALIESQALAHVWFAFRLYYQLTSSL----- 374

```

QY 419 GAADAPGAGAT--AGIHWSQLLYQIGTQWLLDSEALHP 454
Db 375 --WDGSHDCTLQEGVHWYSPFPRYRISLVEPTLLHP 410

RESULT 9
ID 073803 PRELIMINARY; PRT; 442 AA.
AC 073803;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Fugu hedgehog.
GN Fugu hedgehog.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99177347; PubMed=10077531;
RA Gellner K., Brenner S.;
RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
rubripes.";
RL Genome Res. 9:251-258(1999).
DR EMBL; AF056116; AAC34384.1; -.
DR HSP; Q62226; 1VHH.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0016539; P:protein splicing; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR003586; Hedgehog_hintC.
DR InterPro; IPR003587; Hedgehog_hint_N.
DR InterPro; IPR000320; HH signal.
DR InterPro; IPR006141; Intein_S.
DR InterPro; IPR001657; Peptide_C46.
DR Pfam; PF01085; Hint; 1.
DR PRINTS; PR00632; SONICHOOG.
DR ProDom; PD003042; HH signal; 1.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR PROSITE; PS50817; INTEIN_N_TER; 1.
SQ SEQUENCE 442 AA; 49286 MW; 1CB2B423B4B748E8 CRC64;

Query Match 42.5%; Score 1049.5; DB 13; Length 442;
Best Local Similarity 48.5%; Pred. No. 8.1e-70;
Matches 224; Conservative 67; Mismatches 132; Indels 39; Gaps 9;

QY 4 LARCLLVVSSLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASGRVE 63
Db 9 LAQSVLLAANSVWLQGS--CGPGGPGGIRTRPKKAMYKQFFPNLSNNIGASGRAB 66

QY 64 KISRNERKELTPNYPNDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGVKLR 123
Db 67 KXITRNERNELVCNYPNDIVFKDEENTNADRWMTKCKDLNRLALAVNQPGVHLR 126

QY 124 VTEGWDGDGHHSESLHYEGRVADITTSDRDRSKYGMRLARLAVEAGPDWVYVESKAHIHC 183
Db 127 VTEAWDGDGHHPPGSLHYEGRVADITTDRETEKYGLLAQLAVEAGPDWVYVESKYHIHC 186

QY 184 SVKAENSVAAGSGCGPGSATVHLEOGTKLVKDLSPGDRVLAADDGRLLYSDFTFLD 243
Db 187 SVKADHSAVVEKGCDFGWSRVTVAGGFQKSLSLTPODRVMALSETGVVFSPVLLFLH 246

QY 244 RDDGAKVYFVIETREPRERLLTAHLLFVAPNDSATGEPAESGSGPSPGALGPRA 303
Db 247 RDPSSRWRFLSLQTEDGR-RLAVTPHLLVFSDAH-----CGPDSQY---QA 289

```

```

QY 304 LFASRVRPGQRYVVAARDGDRRLPAAVHSVTLSBEAAGAYAPLTAQGTILINRVLASC 363
Db 290 QFASRAQTGTCVLV--HTAGGEVHPSPRIVSIT-EEESVGAYAPLTAAGSVFVDGLVASS 345

QY 364 YAVIEHSWAHRAFAFERLAHALLAALAPARTDRGDSG-----GGDRGGGGGRVA 414
Db 346 YALVEDHQLAHAFGPVRLSSVQLLWABFEESDGSKTPLOPHALVRGDR-----KVC 400

QY 415 L--TAPGAADAPGAGATAGIHWSQLLYQIGTQWLLDSEALHP 454
Db 401 ARNSTSVRSEAGPRGRTSEVHWYQAQLHRLGWIIVNFDLHP 442

RESULT 10
Q9GRA8 PRELIMINARY; PRT; 447 AA.
AC Q9GRA8;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hedgehog protein (Fragment).
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OX NCBI_TaxID=6999;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461145; PubMed=11003837;
RA Niwa N., Inoue Y., Nozawa A., Saito M., Misumi Y., Ohuchi H.,
RA Yoshioka H., Noji S.;
RT "Correlation of diversity of leg morphology in Gryllus bimaculatus
(cricket) with divergence in dpp expression pattern during leg
development.";
RL Development 127:4373-4381(2000).
DR EMBL; AB044709; BAB19658.1; -.
DR HSP; Q62226; 1VHH.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0016539; P:protein splicing; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR003586; Hedgehog_hintC.
DR InterPro; IPR003587; Hedgehog_hint_N.
DR InterPro; IPR000320; HH signal.
DR InterPro; IPR006141; Intein_S.
DR InterPro; IPR001657; Peptide_C46.
DR Pfam; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHOOG.
DR ProDom; PD003042; HH signal; 1.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR PROSITE; PS50817; INTEIN_N_TER; 1.
FT NON_TER 1
SQ SEQUENCE 447 AA; 48008 MW; 29AEFE061C3EE6F0 CRC64;

Query Match 42.2%; Score 1041.5; DB 5; Length 447;
Best Local Similarity 51.4%; Pred. No. 3.3e-69;
Matches 228; Conservative 47; Mismatches 118; Indels 51; Gaps 10;

QY 1 MLLARCLLVVSSLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
Db 15 MLLLRLLWLLLL--LMQGAARACGPGRGARRRRLTPLVFKQHVPNVSEHTLGAAG 71

QY 61 RYEGKISNSERFKELTPNYPNDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120
Db 72 PAERTVADDDPRFDLVNPNYINADIVFKDEEGTGADRLMTQCKDKLNALAI SVMNQPGV 131

QY 121 KLRTVTEGWDGDGHHSESLHYEGRVADITTSDRDRSKYGMRLARLAVEAGFDWVYVESKAH 180

```

```
132 RLAVIEGWDEBEGHAADSLHYEGRAVDVTTSDRRSKYGMARLAVEAGFDWVYETRGH 191
181 IHCSVKAENSVAKSGCGPFSATVHLEQGTGLVKDLSPGR-VLAADDQGLLYSDFL 239
192 IHCSVKSSESQAASGCGFAESTVQ-TPGGLCALAELRVGRGAGGCHGLAFSPVL 250
240 TFLDRDDGAKKVFYVETREPRERLLTLTAHLFLVAPHNDSATGEPEASGGPFGGAL 299
251 LFLDRDP-APRTLLRVRTASGR-TLALTPSHLL-----SVARAGGE----- 290
300 GPALPASRVPGQVYVVAERDGRLLPAAVHSVTLSEEAAGAVAPITAQGTILINRV 359
291 -PEARFADAVPGDALLVAADAGAVR--PDRVLHVAEATRGVAVAPITAAGTVVVDGV 347
360 LASCYAVIEHSHWAHRAFAPFRLAHLAALAPARTDRGDSGGGGRGGRVALTPG 419
348 LASCYAVVGHSLAHSFAPVRAWH-WLTANGHAAPDY----- 384
420 AADAPGAGATAGIHWISQLLYQIG 443
385 AHPPPPFAAPGVHVAKALYSLG 408

RESULT 11
Q7Z264 PRELIMINARY; PRT; 381 AA.
AC Q7Z264;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hedgehog (Fragment).
OS Euscorpius flavicaudis.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Chactoidae; Euscorpiidae; Euscorpius.
OX NCBI_TaxID=100976;
RN [1]
RP SEQUENCE FROM N.A.
RA Simonnet F., Queinnec E.;
RT "Evolution of Arthropods hedgehog gene: expression in Euscorpius
(R Scorpiones) and Artemia (Crustacea).";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY281133; AAP38181.1; -.
FT NON TER 1
SQ SEQUENCE 381 AA; 43047 MW; 0AC2196211C282B6 CRC64;

Query Match 40.9%; Score 1010; DB 5; Length 381;
Best Local Similarity 48.9%; Pred. No. 5.8e-67;
Matches 213; Conservative 54; Mismatches 99; Indels 70; Gaps 8;

QY 31 GKRHPKLTPLAYKQFIPNVAEKTGLGASGRYEGKISRNSERPKELTPNYPDIIFKDEE 90
D 2 GRRQRPKLTPLVFKQHPNVAETTLAASGGQDKINRDETRPKELVPNTNPDIIIFKDEE 61
QY 91 NTGADRLMTQCKDKLNALAI SVMNQPGVKLRVTEGWDEGDGHSSESLHYEGRAVDITT 150
D 62 GTGADRLMTLCREKNTLAI SVMNQPGVKLRVTEGWDEGDGPHAINSLHYEGRAVDITT 121
QY 151 SDRRSKYGMARLAVEAGFDWVYVESKAHHCVSKAENSVAKSGCGPFGSATVHLEQ 210
D 122 SDRRSKYGMARLAVADAGFDWVYVESRAHHCVSSESADAASGCGFGQNSVW-LTKN 180
QY 211 GTKLKDLSPGRVLAADDQGLLYSDFLTFLDDGAKKVFYVETREPRERLLTAAH 270
D 181 GAKKMSLDVNGNEVLDCSRNGFFTSYQVIMFLWDPPKTNROFYVHLNTGORS-ITLTPSH 239
QY 271 LFLVAPHNDSATGEPEASGGPFGGALGPALPASRVPGQVYVVAERDGRRLP 330
D 240 LIFVVGKNT-----RVVETRDVRPGH--YILAKHFRENVTSP 276
QY 331 AVHSVTLSEEAAGAVAPITAQGTILINRVLASCYAVIEHSHWAHRAFAPFLAHA----- 385
D 277 LVTHVT-TIVTGVYAPLTIGNLVVDGVFASCAVINQDQMAHWAFLVPRVADSVKEMV 335
```

```
386 -----LLAALAPARTDRGDSGGGGRGGRVALTPGAADAPGAGATAGIHWYSQ 437
336 FRILEKPHLMERTSPSR-----AP-----DPGIHWYAR 364
438 LLYQIGTWLWLDSEALH 453
365 SLFYIARYVLPSSYLY 380

RESULT 12
Q7Z263 PRELIMINARY; PRT; 421 AA.
AC Q7Z263;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hedgehog (Fragment).
OS Artemia sanfranciscana (Brine shrimp) (Artemia franciscana).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
OX NCBI_TaxID=6661;
RN [1]
RP SEQUENCE FROM N.A.
RA Simonnet F., Queinnec E.;
RT "Evolution of Arthropods hedgehog gene: expression in Euscorpius
(R Scorpiones) and Artemia (Crustacea).";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY281134; AAP38182.1; -.
FT NON TER 1
SQ SEQUENCE 421 AA; 47289 MW; 350B034CFD72817A CRC64;

Query Match 40.3%; Score 995; DB 5; Length 421;
Best Local Similarity 49.0%; Pred. No. 8.7e-66;
Matches 223; Conservative 53; Mismatches 103; Indels 76; Gaps 11;

QY 31 GKRHPKLTPLAYKQFIPNVAEKTGLGASGRYEGKISRNSERPKELTPNYPDIIFKDEE 90
D 2 GRRQRPKLTPLVFKQHPNNSENTLGASGLSEGIKTRDDPRFKDLVPNTNPDIIIFKDEE 61
QY 91 NTGADRLMTQCKDKLNALAI SVMNQPGVKLRVTEGWDEGDGHSSESLHYEGRAVDITT 150
D 62 GTGADRLMTQCKEKLNTLAI SVMNQPGVKLRVTEGWDEGLHASNLSHYEGRAVDITT 121
QY 151 SDRRSKYGMARLAVEAGFDWVYVESKAHHCVSKAENSVAKSGCGPFGSATVHLEQ 210
D 122 SDRRSKYGMARLAVEAGFDWVYVESRAHIHTSVKSSSAARSAGCGFTGDSLVTLDG 181
QY 211 GTKLKDLSPGRVLAADDQGLLYSDFLTFLDRDDGAKKVFYVETREPRERLLTAAH 270
D 182 EQVKMSDVRIGRDLVADVEDGNLIYSEVLLFLDRDVTNROFVKL-TTESGETVELTASH 240
QY 271 LFLVAPHN-----DSATGE---PEASSGGPFGGALGPALPASRVPGQ 313
D 241 LIFTVQSNYEYTMKXLENENIDSSIEETELSDNEGD-----LFVSKL---Q 287
QY 314 RVYVVAER-----DGDRLPAAVHSVTL---SEEA---AGAYAPLTAQGTILINVL 360
D 288 FHTLVAETFAKNVGDYLLVKNKVGKLVQLQVTEVAFSVQTVGVPAPLNTGIIIVNSVA 347
QY 361 ASCYAVIEHSHWAHRAFAP-----FRLAHLAALAPARTDRGDSGGGGRGGRVA 414
D 348 ALLYAVVDVSHLTAHWAFFPLRWYSNFNEAYSILSVASLK-DKNSRTKFFE----- 397
415 LTPAGAADAPGAGATAGIHWISQLLYQIGTWLWLD 449
398 -----GHWYAKLYLDLSHVIVPS 416

RESULT 13
Q8MWG6 PRELIMINARY; PRT; 416 AA.
ID Q8MWG6
AC Q8MWG6;
```







```

RESULT 19
Q9YGV7 PRELIMINARY; PRT; 150 AA.
AC Q9YGV7
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8296;
RN [1]
SEQUENCE FROM N.A.
RA Torok M.A., Izpizua-Belmonte J.C., Gardiner D.M., Bryant S.V.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF031480; AAD18128.1; -.
DR HSP; Q62226; 1VHH.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0007257; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
DR NON_TER 1
DR NON_TER 150
FT SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;
SQ
Query Match 29.7%; Score 733; DB 13; Length 150;
Best Local Similarity 93.3%; Pred. No. 6.6e-47;
Matches 140; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 83 DIIFKDEENTGADRLMTQRCQKLNALAI SVMNOMPVGKLRVTEGWDGHHSESLHYE 142
Db 1 DIIFKDEENTGADRLMTQRCQKLNALAI SVMNOMPVGKLRVTEGWDGHHSESLHYE 60
QY 143 GRAVDITTSDDRSKYGMARLAVEAGPDWVYYSKAHIGSVKAENSVAAKSGCFFGS 202
Db 61 GRAVDITTSDDRSKYGMARLAVEAGPDWVYYSKAHIGSVKAENSVAAKSGCFFAS 120
QY 203 ATVHLEQGQTKLVKDLSPGDRVLAADQGR 232
Db 121 AKVTLEHGVTVPVKDLRPGDRVLAADQGR 150
RESULT 20
Q9XS16 PRELIMINARY; PRT; 139 AA.
AC Q9XS16
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RA Tissue-Touch germ;
RC Koyama E., Iwanoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
RA Bashir M.M., Tucker T., Pacifici M.;
RT "Development of Stratum intermedium and its Role as a Sonic Hedgehog-
RT Signaling Structure During Odontogenesis.";
Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF144100; AAD33926.1; -.
DR HSP; Q62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
DR NON_TER 1
DR NON_TER 139
FT SEQUENCE 139 AA; 15961 MW; C83B0B96982046B1 CRC64;
SQ
Query Match 29.6%; Score 731; DB 6; Length 139;
Best Local Similarity 98.6%; Pred. No. 8.3e-47;
Matches 137; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 43 AYKQFIPNVAEKTIGASGRYEGKISRNSERFKELTPNYPNPDIIIFKDEENTGADRLMTQRC 102
Db 1 AYKQFIPNVAEKTIGASGRYEGKISRNSERFKELTPNYPNPDIIIFKDEENTGADRLMTQRC 60
QY 103 KDKLNALAI SVMNOMPVGKLRVTEGWDGHHSESLHYEGRVDITTSDDRSKYGMARL 162
Db 61 KDKLNALAI SVMNOMPVGKLRVTEGWDGHHSESLHYEGRVDITTSDDRSKYGMARL 120
QY 163 RLAVEAGPDWVYYSKAHIGSVKAENSVAAKSGCFFAS 181
Db 121 RLAVEAGPDWVYYSKAHIGSVKAENSVAAKSGCFFAS 139
RESULT 21
Q9W6C1 PRELIMINARY; PRT; 138 AA.
AC Q9W6C1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sonic hedgehog protein (Fragment).
GN SHH.
OS Eleutherodactylus coqui (Puerto Rican coqui).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Leptodactylidae;
OC Telmatobiinae; Eleutherodactylus.
OX NCBI_TaxID=57860;
RN [1]
SEQUENCE FROM N.A.
RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klymkowsky M.W.,
RA Hanken J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog E. coqui.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF113403; AAD23436.1; -.
DR HSP; Q62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
DR NON_TER 1
DR NON_TER 138
FT SEQUENCE 138 AA; 15751 MW; FF4156A1F74681F0 CRC64;
SQ
Query Match 28.7%; Score 709; DB 13; Length 138;
Best Local Similarity 96.4%; Pred. No. 3.6e-45;
Matches 133; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 48 IPNVAEKTIGASGRYEGKISRNSERFKELTPNYPNPDIIIFKDEENTGADRLMTQRCQKLN 107

```

```
Db 1 IPNVAEKLGAQSGYEGKITRNSRPFKELTPNNSDIIFKDBENTAADRLMTQRCCKLN 60
QY 108 ALAISVWQMPGKLVRTGWDGHHSESLHYEGRAVDITTSDDRSKYGLMLARLAVE 167
Db 61 ALAISVWNLWPGKLVRTGWDGHHSESLHYEGRAVDITTSDDRSKYGLMLARLAVE 120
QY 168 AGFDWYVYSEKAIHCSV 185
Db 121 AGFDWYVYSEKAIHCSV 138

RESULT 22
Q9MW29 PRELIMINARY; PRT; 177 AA.
AC Q9MW29
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Indian hedgehog protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RA Garges P.L., Meyer R.A. Jr., Brown C.A., Price D.K.;
RT "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162914; AAD45372.1; -
DR HSP; G62226; 1VHH.
DR MRO8; C46.003; -
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0016539; P:protein splicing; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR003587; Hedgehog hint_N.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR006141; Intein_S.
DR InterPro; IPR001657; Peptidase_C46.
DR InterPro; IPR001767; Pept_C46_hint.
DR Pfam; PF01085; HH_signal; 1.
DR Pfam; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
DR SMART; SM00306; HintN; 1.
DR PROSITE; PS00817; INTEIN_N_TER; 1.
FT NON_TER 1
FT NON_TER 177
SQ SEQUENCE 177 AA; 19739 MW; CA5871626A565F65 CRC64;

Query Match 28.4%; Score 702; DB 11; Length 177;
Best Local Similarity 74.0%; Pred. No. 1.7e-44;
Matches 131; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

QY 97 LMTQRCCKLNLAISVWQMPGKLVRTGWDGHHSESLHYEGRAVDITTSDDRS 156
Db 1 LMTQRCCKLNLAISVWQMPGKLVRTGWDGHHSESLHYEGRAVDITTSDDRN 60
QY 157 KYGLMLARLAVEAGFDWYVYSEKAIHCSVVAAGCGCTPGSATVHLEQCGTKLVK 216
Db 61 KYGLMLARLAVEAGFDWYVYSEKAIHCSVVAAGCGCTPGSATVHLEQCGTKLVK 120
QY 217 DLSPGDVLAAADGRLLYSDFTFLDRDDGAKVFFVITRPRELLMLTAHLLF 273
Db 121 AVKEDRVLANGEQCNFTFSDFLFLDRBNRLRAFOVIQTQPPRLATPAHLF 177

RESULT 23
Q8MJV7 PRELIMINARY; PRT; 121 AA.
ID Q8MJV7
```

```
AC Q8MJV7;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
GN SSH.
OS Suncus murinus (House shrew) (Musk shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Crocidurinae; Suncus.
OX NCBI_TaxID=9378;
RN [1]
RP SEQUENCE FROM N.A.
RA Ogi H., Tabata M.J., Yamanaka A., Yasui K., Uemura M.;
RT "COMPARISON OF EXPRESSION PATTERNS OF FIBROBLAST GROWTH FACTOR 8, BONE
RT MORPHOGENETIC PROTEIN 4 AND SONIC HEDGEHOG IN JAW DEVELOPMENT OF THE
RL HOUSE SHREW, SUNCUS MURINUS.";
DR EMBL; AB081406; BAC02893.1; -
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13919 MW; D159C4CSDASAEAL6 CRC64;

Query Match 26.3%; Score 649; DB 6; Length 121;
Best Local Similarity 99.2%; Pred. No. 8.7e-41;
Matches 120; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 71 ERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNLAISVWQMPGKLVRTGWD 130
Db 1 ERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNLAISVWQMPGKLVRTGWD 60
QY 131 DGHSESLHYEGRAVDITTSDDRSKYGLMLARLAVEAGFDWYVYSEKAIHCSVKAENS 190
Db 61 DGHSESLHYEGRAVDITTSDDRSKYGLMLARLAVEAGFDWYVYSEKAIHCSVKAENS 120
QY 191 V 191
Db 121 V 121

RESULT 24
O96699 PRELIMINARY; PRT; 185 AA.
ID O96699
AC O96699
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hedgehog protein (Fragment).
GN HH.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX NCBI_TaxID=39708;
RN [1]
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
RT evolution.";
RL Science 0:0-0(1999).
DR EMBL; AF117742; AAD08931.1; -
DR HSP; G62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
```

```
DR InterPro: IPR000320; HH signal.
DR InterPro: IPR001657; Peptidase_C46.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHOG.
DR ProDom: PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAAE662 CRC64;

Query Match 25.1%; Score 619; DB 5; Length 185;
Best Local Similarity 68.2%; Pred. No. 2.7e-38;
Matches 116; Conservative 21; Mismatches 31; Indels 2; Gaps 2;

QY 58 ASGYEGKISNSRRFKELTNYNPDIIFKDEENTGADRLMTQRCCKKLNALAISVMNQW 117
DB 3 ASGPPEGRIITRDEKFRDLVFNYPNPDIFDKDEGTGADRLMTQRCCKKLNTLAISVMNQW 62

QY 118 PGVKLRVTEGWDEGHSEELHVEGRAVDITTSDDRSKYGMRLAVEAGFDWVYVES 177
DB 63 PGVRLRVTEGWDEENSHLDNSLHVEGRAVDITTSDDRSKNGKMLARLAVEAGFDWVFYEN 122

QY 178 KARIHCSVKAENSVAAKSGGCPGASATVHLEQGGTKLVKDLSPGDRVLA 227
DB 123 RSYIHCSVKYTESVGT-CAGCFPSGAVVHTENGPKD-IASLKGNKVLA 170

RESULT 25
Q8QGN4 PRELIMINARY; PRT; 99 AA.
AC Q8QGN4;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21930458; PubMed=11932743;
RA Tanaka M., Munsterberg A., Anderson W.G., Prescott A.R., Hazen N.,
RA Tickle C.;
RT "Fin development in a cartilaginous fish and the origin of vertebrate
RT limbs."
RL Nature 416:527-531 (2002).
DR EMBL; AF393835; AAM0828.1; -.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro: IPR009045; Hedgehog/DD_pept.
DR InterPro: IPR000320; HH_signal.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 99
SQ SEQUENCE 99 AA; 11468 MW; 831864BDB34A1A20 CRC64;

Query Match 21.5%; Score 531; DB 13; Length 99;
Best Local Similarity 98.0%; Pred. No. 4e-32;
Matches 97; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 78 PNYNPDIIFKDEENTGADRLMTQRCCKKLNALAISVMNQWPGVKLRVTEGWDEGHSE 137
DB 1 PNYNPDIIFKDEENTGADRLMTQRCCKKLNSLAISVMNQWPGVKLRVTEGWDEGHSE 60

QY 138 SLHYEGRVDITTSDDRSKYGMRLAVEAGFDWVYVE 176
DB 61 SLHYEGRVDITTSDDRTKYGMRLAVEAGFDWVYVE 99
```

```
RESULT 26
Q8MIN9 PRELIMINARY; PRT; 202 AA.
AC Q8MIN9;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Indian hedgehog (Fragment).
GN IHH.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Semevolos S.A., Nixon A.J.;
RT "Partial nucleotide sequence for equine Indian hedgehog mRNA.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV112896; AAM44071.1; -.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0016539; P:protein splicing; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009045; Hedgehog/DD_pept.
DR InterPro: IPR003586; Hedgehog/hintC.
DR InterPro: IPR003587; Hedgehog/hint_N.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR006141; Intein_S.
DR InterPro: IPR001657; Peptidase_C46.
DR InterPro: IPR001767; Peptidase_C46_hint.
DR InterPro: IPR002375; Pr/PY_rp_transf.
DR Pfam; PF01085; HH_signal; 1.
DR Pfam; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
DR SMART; SM00305; HintC; 1.
DR SMART; PS00817; INTEIN_N_TER; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 21601 MW; 3C9E2872BD3E6F1D CRC64;

Query Match 21.3%; Score 526.5; DB 6; Length 202;
Best Local Similarity 51.1%; Pred. No. 2.4e-31;
Matches 114; Conservative 28; Mismatches 60; Indels 21; Gaps 4;

QY 155 RSKYGMRLAVEAGFDWVYVESKAHICSVKAENSVAAKSGGCPGASATVHLEQGGTKL 214
DB 1 RNKYGLLRLAVEAGFDWVYVESKAHVCSKSEHSAAKTGCCFFAGQVLESARVA 60

QY 215 VKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDGAKKVFYVITREPRERLLLTAHLLPV 274
DB 61 LSAVRPGDRVLANGEDNGNTFTSDVLIFLDRPPDLRAFOVETQDPPRLALTPAHLFT 120

QY 275 APNDSATGEPEASSGPPSGALGFALPASRYEPGQRYVVAERDGRLLPAAVES 334
DB 121 ANNH-----SEPAAHF-----RATFASQVQPGQYVLVA-----GVPGQPARVA 160

QY 335 VTLESEAGAYAPLTAQGITILINRVLASCVAVIEHSHWAHRAF 377
DB 161 VS-THVALGAYAPLTRHGLTVEDVVASCFAAVADHHLAQLAF 202

RESULT 27
Q42128 PRELIMINARY; PRT; 119 AA.
AC Q42128;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Shh (Fragment).
```

GN ME-SHH.  
OS *Oryzias latipes* (Medaka fish) (Japanese ricefish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Belontiiformes; Adrianichthyidae; Oryziinae; *Oryzias*.  
OX NCBI\_TaxID=8090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=bbRR; TISSUE=Embryo;  
RA Araki K.;  
RT "The analysis of medaka haploid development";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB007129; BAA22368.1; -  
DR HSSP; Q62226; 1VHH.  
DR MEROPS; C46.001; -  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0007267; P:cell-cell signaling; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR000320; HH\_signal.  
DR InterPro; IPR001657; Peptidase\_C46.  
DR Pfam; PF01085; HH\_signal; 1.  
DR Pfam; PF01079; Hint; 1.  
DR PRINTS; PR00632; SONICHHOG.  
DR ProDom; PD003042; HH\_signal; 1.  
FT NON\_TER 1  
FT NON\_TER 119  
SQ SEQUENCE 119 AA; 13179 MW; A4868AC2E475CB1B CRC64;  
Query Match 21.2%; Score 523; DB 13; Length 119;  
Best Local Similarity 82.4%; Pred. No. 2.1e-31;  
Matches 98; Conservative 8; Mismatches 13; Indels 0; Gaps 0;  
QY 123 RVTEGWEDGHSESLSHYEGRAVDITTSRDRSKYGLMLARLAVEAGFDWVYVESKAH1H 182  
DB 1 RVTEGWEDGHFEESLHYEGRAVDITTSRDRSKYGLTSLKLAVERAGFDWVYVESKAH1H 60  
QY 183 CSVKAESVAAKSGCGPFGATVHLEGGTKLYKDLSPGRVLAADDQGLLYSDLTFT 241  
DB 61 CSVKAESVAAKSGCGPFGSTVTLNGTORPKVLDQGRVLAADYDGNPNVYTDFTMF 119  
RESULT 28  
ID Q9WUP6 PRELIMINARY; PRT; 129 AA.  
AC Q9WUP6;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
DE Desert hedgehog protein (Fragment).  
GN DHH.  
OS *Rattus norvegicus* (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; *Rattus*.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RA Garges F.L., Meyer K.A. Jr., Brown C.A., Price D.K.;  
RT "Desert hedgehog in the rat";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF148226; AAD31927.1; -  
DR HSSP; Q62226; 1VHH.  
DR GO; GO:0007267; P:cell-cell signaling; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR InterPro; IPR009045; Hedgehog/DD\_pept.  
DR InterPro; IPR000320; HH\_signal.  
DR InterPro; IPR001657; Peptidase\_C46.  
DR Pfam; PF01085; HH\_signal; 1.  
DR PRINTS; PR00632; SONICHHOG.

DR ProDom; PD003042; HH\_signal; 1.  
FT NON\_TER 1  
FT NON\_TER 129  
SQ SEQUENCE 129 AA; 14578 MW; AB33509B3E751319 CRC64;  
Query Match 20.7%; Score 511; DB 11; Length 129;  
Best Local Similarity 71.4%; Pred. No. 1.8e-30;  
Matches 90; Conservative 24; Mismatches 10; Indels 2; Gaps 2;  
QY 23 ACGPGRG-FGRRHPKK-LTPLAYKQFIPNVAETLGCASGRYEGKIGRNSERFKELTPNY 80  
DB 4 SCGPGRGPGVRRYVRKQLVELLYKQVPSMPERTLGCASGPAEGRVTRGSRFRDLVNY 63  
QY 81 NPDIFKDEBNTGADRLMTQCKOKLALALAISSVNVQVGYKLVYTEGWDGDEGHSESLSH 140  
DB 64 NPDIFKDENSGADRLMTERCKERVNALATAVNMMPGVRLRVTEGWDGDEGHHAQDSLH 123  
QY 141 YEGRAV 146  
DB 124 YEGRAV 129  
RESULT 29  
ID Q42234 PRELIMINARY; PRT; 137 AA.  
AC Q42234;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
DE Sonic hedgehog (Fragment).  
GN SHH.  
OS *Coturnix coturnix* (Common quail).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Coturnix.  
OX NCBI\_TaxID=9091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=98167903; PubMed=9435297;  
RA Borycki A.G., Mendham L., Emerson C.P. Jr.;  
RT "Control of sonic patterning by Sonic hedgehog and its downstream  
RT signal response genes";  
RL Development 125:777-790 (1998).  
DR EMBL; AF022882; AAB80949.1; -  
DR MEROPS; C46.002; -  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0007154; P:cell communication; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR003586; Hedgehog\_hintC.  
DR InterPro; IPR001657; Peptidase\_C46.  
DR InterPro; IPR001767; Peptidase\_C46.  
DR Pfam; PF01079; Hint; 1.  
DR PRINTS; PR00632; SONICHHOG.  
DR SMART; SM00305; HintC; 1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 137 AA; 14571 MW; 7FD29DF815AF1532 CRC64;  
Query Match 19.3%; Score 476.5; DB 13; Length 137;  
Best Local Similarity 60.2%; Pred. No. 7.3e-28;  
Matches 103; Conservative 12; Mismatches 21; Indels 35; Gaps 5;  
QY 293 PPSGGLGPRALFASRVFGQVYVVAEDGDRRLPAAVHVSVTLSSEAAGAYAPLTAQG 352  
DB 1 PTSG-----QAFASNVKPGQVYVIGE--GGRQLLPASVSVSREASGAYAPLTAQG 53  
QY 353 TILINRVLASCYAVIEEHSWAHRAFAPEFLAHALALAPARTDRGGSGGGGDRGGGR 412  
DB 54 TILINRVLTSCYAVIEEHSWAHRAFAPEFLAQGLLALCP----- 93  
QY 413 VALTAPGAADAGAG-ATAGIHWYSQLLYQICTWLDSEALHPLGNVYKSS 462  
DB 94 -----DGA--IPFAGTTTGIHWYRLLYHIGSVWLDGDLHPLGNVAPAS 137

```

DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 10017 MW; 0590DB14605D13CE CRC64;

Query Match Similarity 15.9%; Score 392; DB 5; Length 87;
Best Local Similarity 80.5%; Pred. No. 7.5e-22;
Matches 70; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 101 RCKDLNALAISVNMQPGVKLRVTEGWDEDDHGHSESLHYEGRAVDITTSDDRSKYGM 160
Db 1 RCKEKLNLILAVSNMQLRVTEGWDEDDHGHSESLHYEGRAVDITTSDDRSKYGM 60
QY 161 LARLAVEAGFDWVYVESKAIHCSVKA 187
Db 61 LARLAVEAGFDWVYVESKAIHCSVKS 87

RESULT 32
Q91814 PRELIMINARY; PRT; 63 AA.
ID Q91814
AC Q91814;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
GN SHH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RA Wroblewski E.B., Patterson J.M.;
RT "Cloning of sonic hedgehog from the bullfrog Rana catesbeiana.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275308; AAF89204.1; -.
DR HSSP; O62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 63
SQ SEQUENCE 63 AA; 7227 MW; 4C1FC3BD7D6F4B13 CRC64;

Query Match 13.7%; Score 339; DB 13; Length 63;
Best Local Similarity 96.8%; Pred. No. 4.2e-18;
Matches 61; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 125 TEGWDEDDHGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAIHCS 184
Db 1 TEGWDEDDHGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAIHCS 60
QY 185 VKA 187
Db 61 VKA 63

RESULT 33
Q9PRF5 PRELIMINARY; PRT; 64 AA.
ID Q9PRF5
AC Q9PRF5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DR InterPro; IPR000320; HH_signal.

DR O42441 PRELIMINARY; PRT; 80 AA.
AC O42441;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Koester R., Stick R., Loosli F., Wittbrodt J.;
RT "Medaka spalt acts as a target gene of hedgehog signaling.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z97019; CAB09695.1; -.
DR HSSP; O62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9180 MW; 5E18295E6B05A4D7 CRC64;

Query Match 15.9%; Score 392; DB 13; Length 80;
Best Local Similarity 91.2%; Pred. No. 6.7e-22;
Matches 73; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 88 DEENTGADRLMTQCKDKLNLALISVNMQPGVKLRVTEGWDEDDHGHSESLHYEGRAVD 147
Db 1 DEENTGADRLMTQCKDKLNLALISVNMQPGVKLRVTEGWDEDDHGHSESLHYEGRAVD 60
QY 148 ITTSDDRSKYGMRLARLAVE 167
Db 61 ITTSDDRSKYGTLSRLAVE 80

RESULT 31
Q9TX30 PRELIMINARY; PRT; 87 AA.
ID Q9TX30
AC Q9TX30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hedgehog HH protein homolog (Fragment).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95236997; PubMed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
RT "Products, Genetic linkage and limb patterning activity of a murine
hedgehog gene.";
RL Development 120:3339-3353(1994).
DR HSSP; O62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
```

```
DE SONIC hedgehog (Fragment).
GN SHH.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA, and HNI;
RX MEDLINE=20082974; PubMed=10613850;
RA Ohtsuka M., Makino S., Yoda K., Wada H., Naruse K., Mitani H.,
RA Shima A., Ozato K., Kimura M., Inoko H.;
RT "Construction of a linkage map of the Medaka (Oryzias latipes) and
RT mapping of the Da mutant locus defective in dorsoventral patterning.";
RL Genome Res. 9:1277-1287(1999).
DR EMBL; AB030474; BAA85141.1; -.
DR EMBL; AB030473; BAA85140.1; -.
DR HSSP; Q62226; 1VHH.
DR MEROPS; C46.001; -.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 64
FT NON_TER 64
SQ SEQUENCE 64 AA; 7420 MW; 3D4798598D741F08 CRC64;

Query Match 13.4%; Score 332; DB 13; Length 64;
Best Local Similarity 93.8%; Pred. No. 1.4e-17;
Matches 60; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 116 QWPGVKLVTEGWDGHHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVY 175
Db 1 QWPGVKLVTEGWDGHHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVY 60

QY 176 ESKA 179
Db 61 ESKA 64

RESULT 34
Q95ZD8 PRELIMINARY; PRT; 60 AA.
AC Q95ZD8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hedgehog protein (Fragment).
GN HEDGEHOG.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaloulis K., Galliot B.;
RT "Hedgehog in hydra.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011384; CAC50883.1; -.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.

DE SONIC hedgehog (Fragment).
GN SHH.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA, and HNI;
RX MEDLINE=20082974; PubMed=10613850;
RA Ohtsuka M., Makino S., Yoda K., Wada H., Naruse K., Mitani H.,
RA Shima A., Ozato K., Kimura M., Inoko H.;
RT "Construction of a linkage map of the Medaka (Oryzias latipes) and
RT mapping of the Da mutant locus defective in dorsoventral patterning.";
RL Genome Res. 9:1277-1287(1999).
DR EMBL; AB030474; BAA85141.1; -.
DR EMBL; AB030473; BAA85140.1; -.
DR HSSP; Q62226; 1VHH.
DR MEROPS; C46.001; -.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 64
FT NON_TER 64
SQ SEQUENCE 64 AA; 7420 MW; 3D4798598D741F08 CRC64;

Query Match 13.4%; Score 332; DB 13; Length 64;
Best Local Similarity 93.8%; Pred. No. 1.4e-17;
Matches 60; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 116 QWPGVKLVTEGWDGHHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVY 175
Db 1 QWPGVKLVTEGWDGHHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVY 60

QY 176 ESKA 179
Db 61 ESKA 64

RESULT 34
Q95ZD8 PRELIMINARY; PRT; 60 AA.
AC Q95ZD8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hedgehog protein (Fragment).
GN HEDGEHOG.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaloulis K., Galliot B.;
RT "Hedgehog in hydra.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011384; CAC50883.1; -.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
```

```
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6809 MW; 6624F85A015999CC CRC64;

Query Match 12.1%; Score 298; DB 5; Length 60;
Best Local Similarity 90.0%; Pred. No. 4.5e-15;
Matches 54; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 113 VVNQWPGVKLVTEGWDGHHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDW 172
Db 1 VVNQWPGVKLVTEGWDGHHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDW 60

RESULT 35
Q90ZRO PRELIMINARY; PRT; 56 AA.
AC Q90ZRO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Indian hedgehog protein (Fragment).
GN IHH.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kondo S., Naruse K., Shima A.;
RT "Cloning of medaka IHH.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF355590; AAK39706.1; -.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR009045; Hedgehog/DD_pept.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6438 MW; 1304032BACD74327 CRC64;

Query Match 11.7%; Score 289; DB 13; Length 56;
Best Local Similarity 94.6%; Pred. No. 1.9e-14;
Matches 53; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 132 GHHSSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVYYSKAIHCSVKA 187
Db 1 GHHSSESLHYEGRAVDITTSRDRKNKYAMLARLAVEAGFDWVYYSKAIHCSVKS 56

RESULT 36
Q866A4 PRELIMINARY; PRT; 74 AA.
AC Q866A4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative indian hedgehog protein (Fragment).
GN IHH.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=cartilage;
RA Petersen J.P., Amling M., Meenen N.M., Haberland M.;
RT "Gene expression patterns in tissue engineered cartilage.";
```

RESULT 40  
Q9ESH3

```

ID Q9ESH3 PRELIMINARY; PRT; 125 AA.
AC Q9ESH3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Indian hedgehog protein (Fragment).
GN IHH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Stevens D.A., Hasserjian R.P., Robson H., van Noorden S., Siebler T.,
RA Shalet S.M., Williams G.R.;
RT "Hypothyroidism causes growth retardation by disrupting chondrocyte
RT differentiation, matrix deposition and endochondral ossification in
RT epiphyseal growth plate cartilage.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF175209; AAC09197.1; -.
DR MEROPS; C46.003; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0007154; P:cell communication; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003587; Hedgehog_hint_N.
DR InterPro; IPR001657; Peptidase_C46.
DR InterPro; IPR001767; Pept_C46_hint.
DR Pfam; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHOG.
DR SMART; SM00306; HintN; 1.
FT NON_TER 1 125
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 13648 MW; 8FAEA76826DB45FD CRC64;

Query Match 7.6%; Score 187.5; DB 11; Length 125;
Best Local Similarity 34.9%; Pred. No. 2.1e-06;
Matches 5; Conservative 18; Mismatches 42; Indels 35; Gaps 4;

QY 218 LSGDRVLAADDQGRLLYSDFLFLDRDDGAKKVFVYETREPREKLLTAAHLFVAPH 277
Db 7 VKGDRVLAMGEGDGNFTFSDVLIPLDREPRLRAFAQVETQDPRLALTPAHLFIADN 66

QY 278 NDSATGPEASSGSPSGGALGPRALFASRVRPGORVYVVAERDGRLLPAVHSVIL 337
Db 67 H----AEPARF-----RATFARHVFGQYLV-----AGVQASL 98

QY 338 -----SBEAAGAYAFLTAGQTILI 356
Db 99 LGWQPSPTWPLGPMFSRSGTLV 124

RESULT 41
Q9TX32 PRELIMINARY; PRT; 48 AA.
AC Q9TX32;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hedgehog HH protein homolog (fragment).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoida; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
CX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95236597; PubMed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";

RL Development 120:3339-3353(1994).
DR HSSP; O62226; 1VHH.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR003045; Hedgehog/DD_Dept.
DR InterPro; IPR000320; HA_signal.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1 48
FT NON_TER 48 48
SQ SEQUENCE 48 AA; 5347 MW; 98A5F87328613DC6 CRC64;

Query Match 7.1%; Score 175.5; DB 5; Length 48;
Best Local Similarity 73.5%; Pred. No. 4.3e-06;
Matches 36; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 119 GVKLRVTEGWDEGHSESHYEGRAVDITTSDDRSKYGMRLARLAVE 167
Db 1 GIKLRVVEAWDED-OPNVEPLHAEGRAVDITTSDDKNKYGALARLAVE 48

RESULT 42
P91573 PRELIMINARY; PRT; 615 AA.
AC P91573;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ZK377.1 protein.
GN ZK377.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaison N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nhan, M., Hawkins J.;
RT "The sequence of C. elegans cosmid ZK377.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88183; AAB52656.1; -.
DR FIR; T29550; T29550.
DR WormPep; ZK377.1; CE15329.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0007154; P:cell communication; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:001653; P:protein splicing; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003586; Hedgehog_hintC.
```



```

Db 634 VOEFLLTMT-EYKTLRITSRHFMYRNKCGKSPQYIKMLPHD-----GE----- 677
QY 296 GGALGPRALFASVRPQGVVVAERDGRLLPAAVHVSVTLSEEAAGAVAPLTAQTIL 355
Db 678 -----AIFASDLEVGDCVVVLYKGYRQCKIETITRSV-----RTGIYSPLTNNGRII 725
QY 356 INRVLASCYAVIEHS-----WAH 374
Db 726 VNDMLASCYSEIQNTLTQTTFFWAY 750

RESULT 45
Q21835
ID Q21835 PRELIMINARY; PRT; 1137 AA.
AC Q21835;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE C. elegans GRD-1 protein (Corresponding sequence R08B4.1a).
GN GRD-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
QY 189 NSVAAKSGGCPGSGATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDGA 248
Db 923 SALVAATGACPSLDTWV-TTPTGKRMDQIDIGDYVLTADLE-KTYFTPTILWIHREPEK 980
QY 249 KKVFIYIETPRERLLITAAHLF-----VAPHNDSATGPEASGSGPPS 295
Db 981 VOEFLLTMT-EYKTLRITSRHFMYRNKCGKSPQYIKMLPHD-----GE----- 1024
QY 296 GGALGPRALFASVRPQGVVVAERDGRLLPAAVHVSVTLSEEAAGAVAPLTAQTIL 355
Db 1025 -----AIFASDLEVGDCVVVLYKGYRQCKIETITRSV-----RTGIYSPLTNNGRII 1072
QY 356 INRVLASCYAVIEHS-----WAH 374
Db 1073 VNDMLASCYSEIQNTLTQTTFFWAY 1097

RESULT 46
Q7YXC8
ID Q7YXC8 PRELIMINARY; PRT; 1160 AA.
AC Q7YXC8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE C. elegans GRD-1 protein (Corresponding sequence R08B4.1b).
GN GRD-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA White S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68008; CAD91696.1; -.
KW Homeobox.
SQ SEQUENCE 1160 AA; 125411 MW; 77894C7C7B4A4B11 CRC64;

Query Match 5.6%; Score 138.5; DB 5; Length 1160;
Best Local Similarity 24.4%; Pred. No. 0.22;
Matches 50; Conservative 38; Mismatches 68; Indels 49; Gaps 8;

QY 189 NSVAAKSGGCPGSGATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDGA 248
Db 946 SALVAATGACPSLDTWV-TTPTGKRMDQIDIGDYVLTADLE-KTYFTPTILWIHREPEK 1003
QY 249 KKVFIYIETPRERLLITAAHLF-----VAPHNDSATGPEASGSGPPS 295
Db 1004 VOEFLLTMT-EYKTLRITSRHFMYRNKCGKSPQYIKMLPHD-----GE----- 1047
QY 296 GGALGPRALFASVRPQGVVVAERDGRLLPAAVHVSVTLSEEAAGAVAPLTAQTIL 355
Db 1048 -----AIFASDLEVGDCVVVLYKGYRQCKIETITRSV-----RTGIYSPLTNNGRII 1095
QY 356 INRVLASCYAVIEHS-----WAH 374
Db 1096 VNDMLASCYSEIQNTLTQTTFFWAY 1120

RESULT 47
O45992
ID O45992 PRELIMINARY; PRT; 481 AA.
AC O45992;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE C. elegans WRT-7 protein (Corresponding sequence ZK1037.10).
GN WRT-7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Basham V.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81142; CAB03509.2; -.
SQ SEQUENCE 481 AA; 54512 MW; 4CE81DB2D0AAA962 CRC64;

Query Match 5.4%; Score 134.5; DB 5; Length 481;
Best Local Similarity 24.3%; Pred. No. 0.13;
Matches 45; Conservative 34; Mismatches 67; Indels 39; Gaps 6;

```

```

QY 198 CPGSATVHLEOGGKTLVLDSPGDRVLAADDQGRLLYSDFLTFLD-----RDDGAKKVF 252
Db 274 CFPNDVAVNVYKAVKRMDELEIGDWVEALDENG-----EDITFLPVKYLWHRDPQEAE 328
QY 253 YVETREPRELLITAAHLIFVAPHNDSATGEPEASSGSGPPGGALGPALFASVRPG 312
Db 329 FLEFLONGETFTLTKHLV-----TTCRQNS-----SELKISWESISAGKNAG 375
QY 313 QRVVVAERDGRLLPAAVHVTLSL-----EAAGAYAPLTAQGTILINRVLASCYA 365
Db 376 DCFYLAQSE-----ALTKYRLVELDIKRVKTKGIYAPMTSQGHLLVKNKIHSCHS 426
QY 366 VIEEH 370
Db 427 EVDHH 431

RESULT 48
Q9XUV2 PRELIMINARY; PRT; 1021 AA.
AC Q9XUV2;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE K02E2.2 protein.
CN K02E2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81550; CAB04547.1; -.
DR PIR; T23252; T23252.
DR WORMPEP; Q02936; IATO.
DR HSSP; Q02936; IATO.
DR GO; GO:008233; F:peptidase activity; IEA.
DR GO; GO:0007154; P:cell communication; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0016539; P:protein splicing; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007284; DUF398.
DR InterPro; IPR003586; Hedgehog hintC.
DR InterPro; IPR003587; Hedgehog hint N.
DR InterPro; IPR006141; Intein S.
DR InterPro; IPR001767; Pept_Cf6_hint.
DR Pfam; PF04155; DUF398; 4.
DR Pfam; PF01079; Hint; 1.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR PROSITE; P550817; INTEIN_N_TER; 1.
SQ SEQUENCE 1021 AA; 115325 MW; E420668EB232D9B7 CRC64;

Query Match 5.1%; Score 126.5; DB 5; Length 1021;
Best Local Similarity 27.8%; Pred. No. 1.4;
Matches 52; Conservative 34; Mismatches 65; Indels 41; Gaps 10;

QY 198 CPGSATVHLEOGGKTLVLDSPGDRVLAADDQGRLLYSDFLTFLDQDGAKKVYVET 257
Db 802 CFSRDVW-TTPSGKKRMBEIGDVVLTDALK-TALFSAITLWIHREPTVQEFLEIKT 859
QY 258 REPREFLLITAAHLFVA-----PHNDS-----ATGE-----PEASSGSGPPGGAL 299
Db 860 -DNGKTLQTAGHFIVATECRYLPSKNLSLNSTPVRHLIDTLPDSDS----- 907

```

```

QY 300 GPRALFASVRPQGRVYVVAERDGR-RLLPAAVHVTLSLSEAAAYAPLTAQGTILINR 358
Db 908 --ETLASQLKIGECILLI---HNGDQFMQKIDSISKTVS---TGIYSPLTENGRIILVND 959
QY 359 VLASCYAVIEEH 370
Db 960 VLASCYSEVQON 971

RESULT 49
Q76786 PRELIMINARY; PRT; 2639 AA.
AC Q76786;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fibroin.
OS Antheraea pernyi (Chinese oak silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Saturniidae; Saturniinae; Saturniini; Antheraea.
OX NCBI_TaxID=7119;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20496953; PubMed=11040284;
RA Sezutsu H., Yukuhiro K.;
RT "Dynamic rearrangement within the antheraea pernyi silk fibroin gene
RT is associated with four types of repetitive units.";
RL J. Mol. Evol. 51:329-338(2000).
DR EMBL; AF083334; AAC32606.1; -.
DR PIR; T31328; T31328.
SQ SEQUENCE 2639 AA; 216057 MW; 2EE3310DEEB09B9A CRC64;

Query Match 5.0%; Score 123.5; DB 5; Length 2639;
Best Local Similarity 28.5%; Pred. No. 9.2;
Matches 55; Conservative 10; Mismatches 101; Indels 27; Gaps 6;

QY 280 SATGEPEASSGSGPPSGGALGPRALFASVRPQGRVYVVAERDGRRLLPAAVHVTLSL 339
Db 2167 AAAAAAAGSAGGCGGCGGSG---SSAAAAAARAGHERAAGSAA----- 2216
QY 340 EAAGAYAPLTAQGTILINRVLASCVAVIEHSHWRAFPFRLAHALLAALAPARTDRG 399
Db 2217 -AAAAAASGA---GRSGS-----YMGDGGYSDSAAAAAASGAGG 2265
QY 400 DSCGSDRGSGG--GRVALTPGAADAPGAGATAGIHWSQLLYQIGTLLDSEALHPLGM 457
Db 2266 SCGYGGYGYGSDSAAAAAASGAGGAGGCGGYG---GYGSYGSASAAAAA 2321
QY 458 AVKSSXSRGAGGG 470
Db 2322 AAAGSGAGGVGGG 2334

RESULT 50
Q863A2 PRELIMINARY; PRT; 702 AA.
AC Q863A2;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE T-box 2 transcription factor.
CN TBX2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Andelfinger G., Etter L., Dymont M., Hitte C., Galibert F.,
RA Kirkness E., Benson D.W.;
RT "Exclusion of canine Tbx2 and Tbx4 as candidates for tricuspid valve

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 18:16:03 ; Search time 59 seconds  
(without alignments)  
2274.747 Million cell updates/sec

Title: US-09-883-848a-15  
Perfect score: 2469  
Sequence: 1 MLLARCLLVLSLLVCS.....GNVKSXSRGAGGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2467	99.9	475	2 AAR77341	Human son
2	2467	99.9	475	2 AAW94473	Human Shh
3	2467	99.9	475	2 AAY05859	Human Son
4	2467	99.9	475	2 AAW97770	Human Son
5	2467	99.9	475	2 AAY05515	Human Son
6	2467	99.9	475	3 AAY96248	Partial h
7	2467	99.9	475	3 AAY70681	Human Son
8	2467	99.9	475	3 AAY95286	Human Son
9	2467	99.9	475	3 AAY95977	Human Son
10	2467	99.9	475	4 AAB84674	Amino aci
11	2467	99.9	475	4 AAB60265	Human Son
12	2467	99.9	475	4 AAE04687	Human son
13	2467	99.9	475	4 AAE05377	Human Son
14	2467	99.9	475	4 AAG55748	Human son
15	2467	99.9	475	4 AAB31222	Amino aci
16	2467	99.9	475	5 ABB79137	Human son
17	2467	99.9	475	5 AAE14297	Human Son
18	2467	99.9	475	5 AAO20923	Human Shh
19	2467	99.9	475	5 AAU99483	Human son
20	2467	99.9	475	6 ADA26256	Human Son
21	2467	99.9	475	7 ADD25260	Human Son
22	2467	99.9	475	7 ABW00868	Human son
23	2467	99.9	475	7 ADD71383	Human son
24	2407	97.5	462	4 AAB85085	Human son
25	2407	97.5	462	6 ADA09286	Human Son

```

99 1467 59.4 418 2 AAW94472 Zebrafish
100 1467 59.4 418 2 AAY05858 Zebrafish
101 1467 59.4 418 2 AAW97769 Zebrafish
102 1467 59.4 418 2 AAY05514 Zebrafish
103 1467 59.4 418 3 AAY96247 Partial z
104 1467 59.4 418 3 AAY70680 Zebrafish
105 1467 59.4 418 3 AAY95285 Zebrafish
106 1467 59.4 418 3 AAY95976 Zebrafish
107 1467 59.4 418 4 AAB84673 Amino aci
108 1467 59.4 418 4 AAB60264 Zebrafish
109 1467 59.4 418 4 AAE04686 Zebrafish
110 1467 59.4 418 4 AAE05376 Zebrafish
111 1467 59.4 418 4 AAB85737 Zebrafish
112 1467 59.4 418 4 AAG65747 Zebrafish
113 1467 59.4 418 4 AAB31221 Amino aci
114 1467 59.4 418 5 ABB79136 Zebrafish
115 1467 59.4 418 5 AAE14296 Zebrafish
116 1467 59.4 418 5 AAC20322 Zebrafish
117 1467 59.4 418 5 AAU99482 Zebrafish
118 1467 59.4 418 6 ADA26255 Zebrafish
119 1467 59.4 418 7 ADD25259 Zebrafish
120 1467 59.4 418 7 ABW00867 Zebrafish
121 1467 59.4 418 7 ADD71382 Zebrafish
122 1432 58.0 416 2 AAR97658 Zebrafish
123 1432 58.0 416 2 AAW61485 Zebrafish
124 1432 58.0 416 2 AAW94475 Zebrafish
125 1432 58.0 416 2 AAY05862 Zebrafish
126 1432 58.0 416 2 AAW97771 Zebrafish
127 1432 58.0 416 2 AAY05518 Zebrafish
128 1432 58.0 416 3 AAY96251 Zebrafish
129 1432 58.0 416 3 AAY70684 Zebrafish
130 1432 58.0 416 3 AAY95289 Zebrafish
131 1432 58.0 416 3 AAY95980 Zebrafish
132 1432 58.0 416 4 AAB84677 Amino aci
133 1432 58.0 416 4 AAB60268 Zebrafish
134 1432 58.0 416 4 AAE04706 Zebrafish
135 1432 58.0 416 4 AAG65751 Zebrafish
136 1432 58.0 416 5 AAE14300 Zebrafish
137 1432 58.0 416 5 AAO20926 Zebrafish
138 1432 58.0 416 5 AAU99486 Zebrafish
139 1271 51.5 411 2 AAW94474 Human Ind
140 1271 51.5 411 2 AAY05860 Human Ind
141 1271 51.5 411 2 AAW97763 Human Ind
142 1271 51.5 411 2 AAY05516 Human Ind
143 1271 51.5 411 3 AAY96249 Human Ind
144 1271 51.5 411 3 AAY70682 Human Ind
145 1271 51.5 411 3 AAY95287 Human Ind
146 1271 51.5 411 3 AAY95978 Human Ind
147 1271 51.5 411 4 AAB84675 Amino aci
148 1271 51.5 411 4 AAB60266 Human Ind
149 1271 51.5 411 4 AAE04688 Human Ind
150 1271 51.5 411 4 AAG65749 Human Ind

```

## ALIGNMENTS

```

RESULT 1
AAR77341
ID AAR77341 standard; protein; 475 AA.
XX
AC AAR77341;
XX
DT 14-MAR-1996 (first entry)
XX
DE Human sonic hedgehog protein.
XX
KW Human; sonic hedgehog protein; probe; primer; diagnostic;
KW nervous system disorder; gene therapy; antibody.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers

```

```

Peptide 24. 29
FT /note= "conserved sequence (AAR77349)"
FT Misc-difference 463
FT /note= "unspecified amino acid"
XX WO9518856-A1.
XX 13-JUL-1995.
XX 30-DEC-1994; 94WO-US014992.
XX 30-DEC-1993; 93US-00176427.
XX 14-DEC-1994; 94US-00356060.
(HARD ) HARVARD COLLEGE.
(IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
Ingham PW, McMahon AP, Tabin CJ;
WPI; 1995-255060/33.
N-PSDB; AAQ91639.
Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful to
treat degenerative nervous system disorder(s) and in gene therapy.
Claim 17; Page 143-45; 210pp; English.
The sequence represents a human sonic hedgehog protein, homologous to a
Drosophila hedgehog protein (AAR77337), and is encoded by a cDNA isolated
from a human fetal lung cDNA library. Probes and primers derived from the
cDNA hedgehog gene may be used as diagnostic agents for neuromuscular,
autonomic or central nervous system disorders, and the gene may also be
used in gene therapy. Antibodies generated from the protein may be used
as therapeutic or research reagents
XX
SQ Sequence 475 AA;
Query Match 99.9%; Score 2467; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.1e-217; Mismatches 0; Gaps 0;
Matches 475; Conservative 0; Indels 0;
QY 1 MLLARCLLLVLLVSSLLVCSGLACGPGRGFRGRRHPKLTPLAYKQFIPNVAEKTLGASG 60
Db 1 MLLARCLLLVLLVSSLLVCSGLACGPGRGFRGRRHPKLTPLAYKQFIPNVAEKTLGASG 60
QY 61 RYEGKISRNSERFKELTNPYNDIIFKOEENTGADRLMTQRCOKLNALASVNNQWPGV 120
Db 61 RYEGKISRNSERFKELTNPYNDIIFKOEENTGADRLMTQRCOKLNALASVNNQWPGV 120
QY 121 KLRVTEGWDGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
Db 121 KLRVTEGWDGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
QY 181 IHCYKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGRVLAADDQGRLLYSDFLT 240
Db 181 IHCYKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGRVLAADDQGRLLYSDFLT 240
QY 241 FLDRDDGAKKVPYVETREPRERLLLTAAHLIFVAPHNDSATGEPPEASSGGSPGSGALG 300
Db 241 FLDRDDGAKKVPYVETREPRERLLLTAAHLIFVAPHNDSATGEPPEASSGGSPGSGALG 300
QY 301 PRALFASRVPRQRYVYVAERDGRRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360
Db 301 PRALFASRVPRQRYVYVAERDGRRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360
QY 361 ASCYAVIEHSHWAHRAPAPFRLAHALLAALAPARTDRCGDSGGGGRGGRVALTAPGA 420
Db 361 ASCYAVIEHSHWAHRAPAPFRLAHALLAALAPARTDRCGDSGGGGRGGRVALTAPGA 420
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDSALHPLGMVKSXSRGAGGGAREGA 475
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLDSALHPLGMVKSXSRGAGGGAREGA 475

```

RESULT 2  
AAW94473  
ID AAW94473 standard; protein; 475 AA.  
XX AC  
XX AC  
XX AC  
DT 29-APR-1999 (first entry)  
DE Human Shh hedgehog protein sequence.  
XX Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;  
KW brain infarction; cerebral infarction; transient ischaemic attack;  
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;  
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.  
XX OS  
XX Homo sapiens.  
XX FH  
XX Key Location/Qualifiers  
FT Misc-difference 463  
FT /label= unknown  
FT /note= "encoded by NNN"  
XX PN  
XX W09900117-A2.  
XX FD  
XX 07-JAN-1999.  
XX PF  
XX 26-JUN-1998; 98WO-US013387.  
XX PR  
XX 27-JUN-1997; 97US-00883656.  
XX PA  
XX (ONTO-) ONTOGENY INC.  
XX PI  
XX Mahanthappa NK;  
XX DR  
XX WPI; 1999-095458/08.  
XX DR  
XX N-PSDB; AAX16187.  
XX PT  
XX Method for limiting damage to neurons caused by ischaemic or epoxic  
PT conditions - is used for the treatment and prevention of e.g. cerebral  
PT infarction, stroke and transient ischaemic attacks.  
XX PS  
XX Disclosure; Page 72-74; 104pp; English.  
XX A method has been developed for limiting the damage to neuronal cells by  
CC ischaemic or epoxic conditions by administering a ptc (patched)  
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to  
CC neuronal cells can also be limited by administering a gene activation  
CC construct which recombines with the genomic hedgehog gene to provide a  
CC heterologous transcription regulator linked to the coding region of this  
CC gene. Administration of the ptc therapeutic agent is used to protect  
CC cerebral tissues against ischaemic injury; to treat cerebral infarct or  
CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic  
CC attacks. It may also be used as a prophylactic in many other cases of  
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage  
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.  
CC Treatment (which may be prophylactic) is used where ischaemic/epoxic  
CC conditions may cause cerebral hypoxia, or progressive loss of neurons due  
CC to oxygen depletion, including in patients with hypotension. The  
CC treatment reduces CIV by at least 25, particularly at least 70%. The  
CC present sequence represents a hedgehog sequence given in the present  
CC invention  
XX SQ  
XX Sequence 475 AA;  
XX Query Match 99.9%; Score 2467; DB 2; Length 475;  
XX Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
XX Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MLLIARCLLLVSVLLVCSGLACGFGGFGRRHPKXLTPLAYKQIFPNVAEXTLIGASG 60  
DB 1 MLLIARCLLLVSVLLVCSGLACGFGGFGRRHPKXLTPLAYKQIFPNVAEXTLIGASG 60

61 RYEGKISRNSEKLTPTNPNPDIIFKDEBNTGADRLMTORCKDLNALAISVNNQPGV 120  
DB 61 RYEGKISRNSEKLTPTNPNPDIIFKDEBNTGADRLMTORCKDLNALAISVNNQPGV 120  
OY 121 KLRVTEGWDEGHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
DB 121 KLRVTEGWDEGHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
OY 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
OY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGEPASSSGSPSGGALG 300  
DB 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGEPASSSGSPSGGALG 300  
OY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQTILLINRVL 360  
DB 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQTILLINRVL 360  
OY 361 ASCYAVIEHSHWAHRAFPAPFLAHALLAALAPASTDRGDSGGDRGGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPAPFLAHALLAALAPASTDRGDSGGDRGGGGRVALTAPGA 420  
OY 421 ADAPGAGATAGIHWYSQQLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSQQLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGGAREGA 475

RESULT 3  
AAW05859  
ID AAW05859 standard; protein; 475 AA.  
XX AC  
XX AAW05859;  
XX DT  
XX 02-AUG-1999 (first entry)  
XX DE  
XX Human Sonic hedgehog Shh protein.  
XX KW  
XX Sonic hedgehog; Shh; human; epithelial tissue; epithelium;  
KW cutaneous tissue; skin; hair; wound healing; vulnary; burn;  
KW skin grafting; pressure sore; ulcer; ulcerative colitis; alopecia;  
KW psoriasis; keratosis; acne; comedogenic lesion; folliculitis;  
KW pseudofolliculitis; keratocanthoma; callosities; Darier's disease; scar;  
KW autoimmune disease; pemphigus; epidermolysis; lupus lesion;  
KW desquamative lesion; carcinoma; therapy; hedgehog therapeutic;  
KW ptc therapeutic; patched.  
XX OS  
XX Homo sapiens.  
XX FH  
XX Key Location/Qualifiers  
FT Misc-difference 463  
FT /note= "encoded by NNN"  
XX PN  
XX W09920298-A1.  
XX PD  
XX 29-APR-1999.  
XX PF  
XX 20-OCT-1998; 98WO-US022227.  
XX PR  
XX 20-OCT-1997; 97US-00955552.  
XX PR  
XX 11-SEP-1998; 98US-00151999.  
XX PA  
XX (ONTO-) ONTOGENY INC.  
XX PI  
XX Wang EA;  
XX WPI; 1999-288170/24.  
XX DR  
XX N-PSDB; AAX25622.  
XX PT  
XX Use of hedgehog polypeptides on patched therapeutics.  
XX PS  
XX Claim 26; Page 127-128; 146pp; English.

XX The present sequence represents human Sonic hedgehog protein Shh. The  
CC invention relates to a method for modulating the growth state an  
CC epithelial cell by ectopically contacting the epithelial cell, in vitro  
CC or in vivo, with a hedgehog therapeutic (i.e. a hedgehog polypeptide or  
CC gene therapy construct) or ptc therapeutic (i.e. a small organic molecule  
CC that mimics the effect of hedgehog proteins on patched signalling, or  
CC activates or potentiates patched signalling) in an amount effective to  
CC alter the rate of proliferation of the epithelial cell. The hedgehog  
CC therapeutic preferably comprises at least a bioactive extracellular  
CC portion of a hedgehog protein (see AA05851-52) encoded by a vertebrate  
CC hedgehog gene (see AA25617-25), especially a human hedgehog gene.  
CC Promotion of proliferation of epithelial cells can be used to control a  
CC wound healing process in e.g. burn treatment, skin regeneration, skin  
CC grafting, pressure sore treatment, dermal ulcer treatment, post surgery  
CC scar reduction or treatment of ulcerative colitis (claimed). It can also  
CC be used to induce hair growth for the treatment of alopecia (claimed).  
CC Inhibition of the growth of epithelial tissue can be used to treat or  
CC prevent hyperplastic or neoplastic conditions, e.g. psoriasis, keratosis,  
CC acne, comedogenic lesions, folliculitis and pseudofolliculitis,  
CC keratoacanthoma, callosities, Darier's disease, keloids, hypertrophic  
CC scars, or autoimmune disorders, e.g. aphthous ulcers, pemphigus vulgaris,  
CC pemphigus foliaceus, pemphigus vegetans, pemphigus erythematosus,  
CC epidermolysis, lupus lesions, degenerative lesions or carcinomas. The  
CC methods can also be used to counteract the effects of ageing on skin  
XX SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 2; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLARCLLLVSVLLVCSGLACGPGRGKRPKLTPLAYKQFIENVAEKTLCASG 60  
DB 1 MLLARCLLLVSVLLVCSGLACGPGRGKRPKLTPLAYKQFIENVAEKTLCASG 60  
QY 61 RYEGKISRNSERFKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
DB 61 RYEGKISRNSERFKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
QY 121 KLRTVEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180  
DB 121 KLRTVEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180  
QY 181 IHCSVKAENSVAKSGCFFGSAIVHLEQGGTKLVKDLSPGRVLAAADQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGCFFGSAIVHLEQGGTKLVKDLSPGRVLAAADQGRLLYSDFLT 240  
QY 241 FLRDDGAKKVFYVETREPRERLLLTAHLLFVAPHNDSATGEPAASSGSGPPSGGALG 300  
DB 241 FLRDDGAKKVFYVETREPRERLLLTAHLLFVAPHNDSATGEPAASSGSGPPSGGALG 300  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASRVPRGQVYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCVAVIEEHSWAHRAPFRLAHALLAALAPARTDRGSDGSGDGGGRVALTAPCA 420  
DB 361 ASCVAVIEEHSWAHRAPFRLAHALLAALAPARTDRGSDGSGDGGGRVALTAPCA 420  
QY 421 ADAPGAGATAGIHWYSOLLVQIGTWLLDSBALHPLGNVAKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSOLLVQIGTWLLDSBALHPLGNVAKSSXSRGAGGAREGA 475

RESULT 4  
AAW97770  
ID AAW97770 standard; protein; 475 AA.

XX AAW97770;

AC AAW97770;

XX 21-MAY-1999 (first entry)

DT 21-MAY-1999 (first entry)

XX

DE Human Sonic hedgehog (Shh) protein.  
XX Sonic hedgehog; Shh protein; human; dopaminergic; GABA-nergic;  
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;  
KW Huntington's disease; amyotrophic lateral sclerosis; cerebral ischaemia;  
XX hypoxia; neuroprotective; therapy.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Misc-difference 463 /note= "encoded by NNN"  
FT WO9904775-A2.  
XX 04-FEB-1999.  
XX 24-JUL-1998; 98WO-US015419.  
XX 24-JUL-1997; 97US-00900220.  
XX (ONTO-) ONTOGENY INC.  
XX Misao N, Wang M, Mahanthappa NK, Pang X;  
XX WPI; 1999-142578/12.  
XX N-PSDB; AAX07276.  
XX Increasing the survival of neuronal, dopaminergic and GABA-nergic cells -  
PT by using a ptc therapeutic such as a protein kinase inhibitor, or an  
PT agent derived from hedgehog polypeptides, useful in the treatment of  
PT Parkinson's disease.  
XX Disclosure; Page 93-95; 138pp; English.  
XX This polypeptide is human Shh Sonic hedgehog protein. The invention is  
CC based on the finding that hedgehog proteins are useful as protective  
CC agents in the treatment and prophylaxis of neurodegenerative disorders  
CC resulting from the loss of dopaminergic and/or GABA-nergic neurons, or  
CC the general loss of tissue from the substantia nigra. Exemplary disorders  
CC include Parkinson's disease, Huntington's disease (both claimed),  
CC amyotrophic lateral sclerosis and cerebral ischaemia. The invention  
CC relates to hedgehog therapeutics (i.e. hedgehog polypeptides and gene  
CC therapy constructs e.g. constructs encoding recombinant hedgehog  
CC polypeptides and trans-activation constructs for altering hedgehog gene  
CC regulatory sequences) and ptc therapeutics (i.e. agents which mimic the  
CC effect of naturally occurring hedgehog proteins on patched signalling)  
CC that are effective in both human and animal subjects. Human Ihh and Dhh  
CC polypeptides (see AAW97763-64) are preferred. The products can also be  
CC used for the maintenance of differentiated neurons in cultures, and to  
CC enhance the implantation of such neuronal cells in an animal. They can be  
CC used to prevent or treat neurodegenerative conditions arising from the  
CC use of certain drugs, and in the prevention and/or treatment of hypoxia,  
CC e.g. as a neuroprotective agent  
XX SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 2; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSVLLVCSGLACGPGRGKRPKLTPLAYKQFIENVAEKTLCASG 60  
DB 1 MLLARCLLLVSVLLVCSGLACGPGRGKRPKLTPLAYKQFIENVAEKTLCASG 60  
QY 61 RYEGKISRNSERFKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
DB 61 RYEGKISRNSERFKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
QY 121 KLRTVEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180  
DB 121 KLRTVEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180

QY 181 IHCSVKAENSVAAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDGGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDGGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVYIETREPRERLLTAAHLFFVAPHNDGATGEPEASSGSGPPSGGALG 300  
DB 241 FLDRDDGAKKVFYVYIETREPRERLLTAAHLFFVAPHNDGATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVTLSERAAAGAYAPLTAQGTILLINRVL 360  
DB 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVTLSERAAAGAYAPLTAQGTILLINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPFRLAHLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGNVAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGNVAVKSSXSRGAGGAREGA 475

RESULT 5

AA05515  
ID AAY05515 standard; protein; 475 AA.

AC AAY05515;

DT 05-JUL-1999 (first entry)

XX Human Sonic hedgehog protein Shh.

XX Sonic hedgehog; Shh protein; human; hedgehog therapeutic;  
KW ptc therapeutic; patched; signal transduction; muscle atrophy; cachexia;  
KW muscular myopathy; myoblastic sarcoma; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 463 /note= "encoded by NNN"

XX WO9910004-A2.

XX 04-WAR-1999.

XX 28-AUG-1998; 98WO-US017922.

XX 29-AUG-1997; 97US-0057394P.

XX (ONTO-) ONTOGENY INC.

XX Bladgen CS, Currie PD, Ingham PW, Hughes SM;

XX WPI; 1999-243557/20.

XX N-PSDB; AAX25103.

XX A new method to regulate muscle growth.

XX Disclosure; Page 118-120; 130pp; English.

XX The present sequence is human Sonic hedgehog protein Shh. The invention  
relates to a method for modulating the formation and/or maintenance of  
muscle tissue by ectopically contacting muscle cells, especially muscle  
stem/progenitor cells, in vitro or in vivo, with a hedgehog therapeutic  
(i.e. hedgehog polypeptides and gene therapy constructs) or ptc  
therapeutic (i.e. a small organic molecule that mimics the effect of  
hedgehog proteins on patched signalling, or activates or potentiates  
patched signalling). In an amount effective to alter the growth state of  
the treated cells. Also claimed is a method for treatment or prevention  
of disorders of, or surgical or cosmetic repair of, such muscle tissues,  
by administering a hedgehog polypeptide or ptc therapeutic. The disorder  
may be muscle atrophy, in particular skeletal muscle atrophy or cardiac  
muscle atrophy, cachexia, or muscular myopathy (all claimed). The

CC hedgehog polypeptide or ptc therapeutic can inhibit growth of myoblastic-  
derived tissue to provide treatment of hyperplastic or neoplastic growth  
of muscle tissue such as in myoblastic sarcoma (also claimed). The  
CC hedgehog therapeutic preferably comprises at least a bioactive  
extracellular portion of a hedgehog protein (see AAY05510-19) encoded by  
CC a vertebrate hedgehog gene (see AAX25098-107), especially a human  
CC hedgehog gene

XX SQ Sequence 475 AA;

Query Watch 99.9%; Score 2467; DB 2; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MILLARCLLLVSSLLVCSGLACGPGFGKRRHPKLLTPLAYQFIPNVAEKLIGASG 60

DB 1 MILLARCLLLVSSLLVCSGLACGPGFGKRRHPKLLTPLAYQFIPNVAEKLIGASG 60

QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAISSVMQWPGV 120

DB 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAISSVMQWPGV 120

QY 121 KLRVTEGWDEDDGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDVWYYESKAH 180

DB 121 KLRVTEGWDEDDGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDVWYYESKAH 180

QY 181 IHCSVKAENSVAAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDGGRLLYSDFLT 240

DB 181 IHCSVKAENSVAAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDGGRLLYSDFLT 240

QY 241 FLDRDDGAKKVFYVYIETREPRERLLTAAHLFFVAPHNDGATGEPEASSGSGPPSGGALG 300

DB 241 FLDRDDGAKKVFYVYIETREPRERLLTAAHLFFVAPHNDGATGEPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVTLSERAAAGAYAPLTAQGTILLINRVL 360

DB 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVTLSERAAAGAYAPLTAQGTILLINRVL 360

QY 361 ASCYAVIEHSHWAHRAFPFRLAHLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420

DB 361 ASCYAVIEHSHWAHRAFPFRLAHLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420

QY 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGNVAVKSSXSRGAGGAREGA 475

DB 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGNVAVKSSXSRGAGGAREGA 475

RESULT 6

AA056248  
ID AAY96248 standard; protein; 475 AA.

XX AAY96248;

XX 11-SEP-2000 (first entry)

XX Partial human Shh.

XX Human; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;  
KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;  
KW chronic inflammatory demyelinating polyneuropathy; CIPD; gene therapy;  
KW infection; inflammation; hereditary neuropathy;  
KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;  
KW multiple myeloma; nutritional imbalance; kidney disease;  
KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;  
KW Tangier disease; Krabbe's disease; Metachromatic leukodystrophy;  
KW Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;  
KW Amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;  
KW Hereditary sensory neuropathy Type II; HSN II; B-cell lymphoma;  
KW Waldenström's Macroglobulinaemia; Chronic Lymphocytic Leukaemia;  
KW neuroprotective; cytoprotective; patched-mediated signal transduction.  
XX Homo sapiens.

FH Key Location/Qualifiers  
 FT Misc-difference 463  
 FT /label= unknown  
 FT /note= "Encoded by NNN"  
 XX WO200027422-A2.  
 XX 18-MAY-2000.  
 XX 08-NOV-1999; 99WO-US026334.  
 XX 06-NOV-1998; 98US-00187387.  
 XX (BIOJ) BIOGEN INC.  
 XX (ONTO-) ONTOGENY INC.  
 XX Galdes A, Mahanthappa N;  
 XX WPI; 2000-387341/33.  
 DR N-PSDB; AAA30279.  
 XX Novel method of preventing deterioration of peripheral nerves, useful for  
 PT treating or preventing neuropathy, e.g. where associated with diabetes or  
 PT viral infection, by administering hedgehog or patched agent.  
 XX Claim 7; Page 134-135; 152pp; English.  
 XX The present sequence is the partial human sonic hedgehog protein, Shh.  
 CC This sequence inhibits expression of the patched gene which has been  
 CC implicated in neuromuscular disorders (neuropathies). This sequence may  
 CC therefore be used for treating neuromuscular disorders i.e. preventing  
 CC degradation in function of motor or sensory nerves and protecting  
 CC peripheral nerve cells under conditions that normally cause neuropathy. A  
 CC variety of neuromuscular disorders may be treated: Guillain-Barre  
 CC syndrome, GBS; peripheral neuropathy; diabetic neuropathy; alcohol-  
 CC induced neuropathy; chronic inflammatory demyelinating polyneuropathy,  
 CC CPD; infection-induced neuropathy, including HIV infection; inflammation  
 CC -induced neuropathy; hereditary neuropathy e.g. Charcot-Marie-Tooth  
 CC disease (CMT); Familial Amyloidotic neuropathy, Refsum's disease,  
 CC Abetalipoproteinemia, Tangier disease, Krabbe's disease, Metachromatic  
 CC leukodystrophy, Fabry's disease, Dejerine-Sottas syndrome, Hereditary  
 CC sensory neuropathy Type II (HSN II) and Amyotrophic lateral sclerosis  
 CC (ALS); acute neuropathy e.g. Miller-Fisher syndrome; neuropathy caused by  
 CC vasculitis; neuropathy associated with tumors e.g. lung cancer, multiple  
 CC myeloma, B-cell lymphoma, Waldenstrom's Macroglobulinemia, Chronic  
 CC Lymphocytic Leukemia; neuropathy associated with: amyloidosis,  
 CC nutritional imbalance, kidney disease, trauma; and hypothyroid  
 CC neuropathy. The coding sequence may be used in gene therapy of the above  
 CC disorders  
 XX SQ Sequence 475 AA;  
 Query Match 99.98; Score 2467; DB 3; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGASG 60  
 DB 1 MLLARCLLLVSSLLVCSGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGASG 60  
 QY 61 RYEGKISRNSEKFKELTPNYPDII FKDEENTGADRLMTORCKDKNALAISVMNQPGV 120  
 DB 61 RYEGKISRNSEKFKELTPNYPDII FKDEENTGADRLMTORCKDKNALAISVMNQPGV 120  
 QY 121 KLRVTEGWDEGDHSESLHYEGRVADITTSDDRSKYGMARLAVEAGFDWYVESKAH 180  
 DB 121 KLRVTEGWDEGDHSESLHYEGRVADITTSDDRSKYGMARLAVEAGFDWYVESKAH 180  
 QY 181 IHCSVKAENSVAKSGCGCFPGSATVHLEQGTGLVKDLSFGDRVLADDDGRLLYSDFLT 240  
 DB 181 IHCSVKAENSVAKSGCGCFPGSATVHLEQGTGLVKDLSFGDRVLADDDGRLLYSDFLT 240  
 QY 241 FLDRDDGAKKVFYIETREPRERLLTAALLFVAPHNDSATGEPASSGSGPPSGGALG 300

DB 241 FLDRDDGAKKVFYIETREPRERLLTAALLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 QY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
 DB 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
 QY 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDRGGSDGGGDRGGGRVALTAPGA 420  
 DB 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDRGGSDGGGDRGGGRVALTAPGA 420  
 QY 421 ADAPCAGATAGIHWSQLLYQIGTWLLDSBALHPLGMVXSSXSRGAGGAREGA 475  
 DB 421 ADAPCAGATAGIHWSQLLYQIGTWLLDSBALHPLGMVXSSXSRGAGGAREGA 475  
 RESULT 7  
 AAY70681  
 ID AAY70681 standard; protein; 475 AA.  
 XX AC AAY70681;  
 XX 18-JUL-2000 (first entry)  
 XX Human Sonic hedgehog (Shh) protein.  
 KW Sonic hedgehog; Shh; human; growth modulator; therapeutic agent; lung;  
 KW hedgehog; hh; patched; ptc; fibroblast growth factor; fgf-10;  
 KW antiproliferative; anticancer; vulnary; antineumatic; hypotensive;  
 KW anti-inflammatory; antiallergic; antiasthmatic; antiarthritic; tuberculostatic; asthma;  
 KW antimicrobial; antiallergy; treatment; prevention; lung diseases; cancer;  
 KW cystic fibrosis; bronchitis; emphysema; respiratory distress syndrome;  
 KW tuberculosis; wound healing; lung transplantation.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Misc-difference 463  
 FT /label= Unknown  
 FT /note= "Encoded by NNN"  
 XX WO200015246-A2.  
 XX 23-MAR-2000.  
 XX 10-SEP-1999; 99WO-US020500.  
 XX 11-SEP-1998; 98US-0099952P.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Pepicelli C, Lewis P, McMahon AP;  
 XX WPI; 2000-271252/23.  
 XX N-PSDB; AAZ52262.  
 XX Modulation of lung tissue or cell growth rate used for treating or  
 PT preventing damage to lung tissue comprises ectopically contacting tissue  
 PT with hedgehog therapeutic, patched therapeutic or fibroblast growth  
 PT factor-10.  
 XX Claim 14; Page 124-127; 143pp; English.  
 CC The patent discloses a method for modulating the growth state of  
 CC epithelial or mesenchymal cells of the lung, by ectopically contacting  
 CC the tissue with a therapeutic agent, that can effectively alter the rate  
 CC of proliferation of cells. This agent can be selected from hedgehog (hh),  
 CC patched (ptc) or fibroblast growth factor (fgf)-10 therapeutics. It  
 CC involves a direct or indirect antagonism of patched-mediated regulation  
 CC of gene expression. This method is useful for the treatment or prevention  
 CC of lung diseases, like cancer, cystic fibrosis, bronchopneumoconiosis,  
 CC bronchitis, bronchopneum, sarcoidosis, silicosis, eosinophilic granuloma,  
 CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress

CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary  
 CC pulmonary hypertension. It is also used to control wound healing or other  
 CC reformatory processes in the lung and augment lung transplantation. The  
 CC present sequence is the human sonic hedgehog (Shh) protein, essential for  
 CC development of the respiratory system. Hedgehog polypeptides can be used  
 CC to control the formation and/or maintenance of the lung tissue  
 XX  
 SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 3; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Qy 121 KLRTGDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
 Db 121 KLRTGDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
 Qy 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
 Qy 241 FLDRDDGAKKVFYVIETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFYVIETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Qy 301 PRALFASVRPGQRYVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASVRPGQRYVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
 Qy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
 Qy 421 ADAPGAGATAGIHWSYQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGGAREGA 475  
 Db 421 ADAPGAGATAGIHWSYQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGGAREGA 475

RESULT 8  
 AA95286  
 ID AA95286 standard; protein; 475 AA.  
 XX  
 AC AA95286;

DT 12-SEP-2000 (first entry)

DE Human Sonic hedgehog Shh protein.

XX Sonic hedgehog; Shh; human; excitotoxicity; Parkinson's disease;  
 KW Huntington's disease; neuronal degeneration; neuroprotective;  
 KW dopaminergic; GABAergic; substantia nigra; therapy.  
 XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 463  
 FT /note= "encoded by NNN"

XX WO200035948-A1.

XX 22-JUN-2000.

XX 03-DEC-1999; 98WO-US028721.

XX 03-DEC-1998; 98WO-US025676.

PR 27-JAN-1999; 98US-00238243.  
 PR 03-JUN-1999; 99US-00325602.  
 XX (BIOJ) BIOGEN INC.  
 PA (ONTO-) ONTOGENY INC.  
 XX  
 PI Galdes A, Mahanthappa N;

DR WPI; 2000-431570/37.  
 DR N-PSDB; AAA27881.

XX Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,  
 PT senile dementia and Korsakoff's disease, by using lipophilic modified  
 PT hedgehog polypeptide.

PS Disclosure; Page 157-159; 174pp; English.

XX The present sequence of that of human Sonic hedgehog (Shh) protein. The  
 CC invention relates to a method for promoting the survival and/or  
 CC functional performance of neuronal cells, especially substantia nigra,  
 CC dopaminergic or GABAergic neurons that are susceptible to exotoxicity, by  
 CC contacting the cells, in vitro or in vivo, with a lipophilic (e.g.  
 CC cholesterol) modified hedgehog polypeptide. The method is used to treat  
 CC or prevent Parkinson's disease, Huntington's disease, diencephalic  
 CC poisoning, spinal cord trauma, hypoglycemia, mechanical trauma to the  
 CC nervous system, senile dementia, Korsakoff's disease, schizophrenia, AIDS  
 CC dementia, multi-infarct dementia, mood disorders, depression, chemical  
 CC toxicity, neuronal damage associated with uncontrolled seizures such as  
 CC epileptic seizures, neuronal injury associated with HIV and AIDS,  
 CC neurodegeneration associated with Down's syndrome, neuropathic pain  
 CC syndrome, olivopontocerebral atrophy, amyotrophic lateral sclerosis,  
 CC mitochondrial abnormalities, Alzheimer's disease, hepatic encephalopathy,  
 CC Tourette's syndrome and drug addiction (all claimed). The lipophilic  
 CC modified hedgehog polypeptide is also useful for promoting survival  
 CC and/or functional performance of neuronal cells susceptible to  
 CC exotoxicity  
 XX

SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 3; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Qy 121 KLRTGDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
 Db 121 KLRTGDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
 Qy 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
 Qy 241 FLDRDDGAKKVFYVIETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFYVIETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Qy 301 PRALFASVRPGQRYVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASVRPGQRYVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
 Qy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
 Qy 421 ADAPGAGATAGIHWSYQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGGAREGA 475

Db 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSSXRGAGGAREGA 475

RESULT 9  
AAV95977  
ID AAV95977 standard; protein; 475 AA.  
AC AAV95977;  
XX  
XX 05-DEC-2000 (first entry)  
DT  
XX Human Sonic hedgehog Shh protein.  
DE  
XX Sonic hedgehog; Shh; human; agonist; antagonist; lipid modulator;  
KW vacuole; cholesterol disorder; lipid disorder; lipid metabolism;  
KW lipid storage; lipid transport; apolipoprotein; triglyceride;  
KW hypercholesterolemia; abetalipoproteinemia; hypobetalipoproteinemia;  
KW chylomicron retention; Anderson's disease; fat absorption;  
KW atherosclerosis; obesity; weight loss; vitamin A disorder;  
KW vitamin E disorder; antilipemia; anorectic; antiarteriosclerotic;  
KW gene therapy; diagnosis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 463  
FT /note= "encoded by NNN"  
FT  
FT WO200051628-A2.  
PN  
XX  
XX 08-SEP-2000.  
PD  
XX 03-MAR-2000; 2000WO-US0055662.  
XX  
XX 03-MAR-1999; 99US-0122640P.  
PR  
XX 15-MAR-1999; 99US-0124446P.  
XX  
XX (BIO ) BIOGEN INC.  
PA  
XX  
XX Burkly L, Wang LC;  
PI  
XX WPI; 2000-611340/58.  
XX  
XX N-PSDB; AAA95977.  
DR  
XX  
XX  
XX  
XX  
XX Use of lipid modulators (e.g. hedgehog agonists or antagonists) for  
PT modulating lipid metabolism and storage, especially useful for treating  
FT lipid metabolism or cholesterol disorders, e.g. obesity or  
FT hypercholesterolemia.  
XX  
XX  
XX Disclosure; Page 118-120; 136pp; English.  
PS  
XX  
XX The present sequence of that of human Sonic hedgehog (Shh) protein. The  
CC invention provides claimed methods for modulating lipid metabolism, for  
CC modulating vacuole formation in intestinal epithelial cells, for  
CC modulating the accumulation of fat in intestinal epithelial cells, for  
CC treating a cholesterol disorder and for treating a lipid metabolism  
CC disorder in an animal (especially a human) by administering a lipid  
CC modulator selected from a hedgehog antagonist or hedgehog agonist. In  
CC particular, the lipid metabolism disorder is a lipid storage disorder, a  
CC lipid transport disorder, an apolipoprotein disorder, a triglyceride  
CC disorder, e.g. a triglyceride metabolism disorder, a triglyceride  
CC transport disorder or a triglyceride storage disorder; a diet-induced  
CC hypercholesterolemia, hypercholesterolemia, abetalipoproteinemia,  
CC hypobetalipoproteinemia; a chylomicron-retention disorder, Anderson's  
CC disease, a fat absorption disorder, e.g. obesity or associated with  
CC weight loss, normotriglyceridemic abetalipoproteinemia, an apolipoprotein  
CC -B100 deficiency, a fat soluble vitamin disorder where the fat soluble  
CC vitamin is vitamin A or E, or atherosclerosis (all claimed). The hedgehog  
CC antagonist binds to the hedgehog receptor, but does not elicit a  
CC response. It is preferably a hedgehog mimetic, a modified hedgehog  
CC protein, e.g. an inactive hedgehog variant, or an anti-hedgehog  
CC homologue, especially a human, chimeric or humanised antibody. The  
CC methods are useful in preventing these disorders or protecting a subject

CC from these disorders. The hedgehog antagonist and agonist are also useful  
CC in diagnosis and research associated with these disorders. The lipid  
CC modulators may also be used as a part of a gene therapy protocol to  
CC deliver polynucleotides encoding these lipid modulators  
XX  
XX Sequence 475 AA;  
SQ  
Query Match 99.9%; Score 2467; DB 3; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2,1e-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFGKRRHPKLLTPLAYKQFIPNVAEKTLCASG 60  
Db 1 MLLARCLLLVSSLLVCSGLACGPGRGFGKRRHPKLLTPLAYKQFIPNVAEKTLCASG 60  
QY 61 RYEGKISRNSERFKELTPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120  
QY 121 KLRVTEGWDEGDHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
Db 121 KLRVTEGWDEGDHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
QY 181 IHCSVKAENSVAAGSGGCGFGSATVHLEOGSTKLVDLSPGDRVLAADDOGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCGFGSATVHLEOGSTKLVDLSPGDRVLAADDOGRLLYSDFLT 240  
QY 241 FLDRDDGAKVYVIETREPRERLLTAHLLFVAPHNDSATGEPEASSSGSPSGGALG 300  
Db 241 FLDRDDGAKVYVIETREPRERLLTAHLLFVAPHNDSATGEPEASSSGSPSGGALG 300  
QY 301 PRALFASRVREGQRYVVAERDGRLLPAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVREGQRYVVAERDGRLLPAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAFPAPFLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGA 420  
Db 361 ASCYAVIEEHSWAHRAFPAPFLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSSXRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSSXRGAGGAREGA 475  
RESULT 10  
AAB84674  
ID AAB84674 standard; protein; 475 AA.  
XX  
XX AAB84674;  
XX  
XX 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of a human hedgehog (Shh) polypeptide.  
DE  
XX Shh gene; hedgehog gene; T lymphocyte; patched gene; infection; diabetes;  
KW nutritional deficiency; graft rejection; hyperacute response;  
KW cornua transplant; autoimmune disorder; multiple sclerosis; psoriasis;  
KW atopic dermatitis; inflammatory disease; proliferative disease;  
KW hyperproliferative disease; eczematous dermatitis; urticaria; vasculitis;  
KW scleroderma.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH Misc-difference 463  
FT /note= "unspecified residue encoded by NNN"  
FT  
XX WO200140438-A2.  
XX  
XX 07-JUN-2001.  
PD  
XX 30-NOV-2000; 2000WO-US032590.  
PF  
XX

PR 30-NOV-1999; 99US-0168112P.  
XX (CURI-) CURIS INC.  
XX Crompton T;  
XX WPI; 2001-441484/47.  
DR N-PSDB; AAH28451.  
XX Modulating immune function comprises administration of a hedgehog or ptc  
PT agent, for treating e.g. diabetes, eczematous dermatitis, urticaria or  
XX vasculitis.  
PS Claim 4; Page 84-86; 105pp; English.  
XX The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog  
CC gene products and signal transduction pathways involving hedgehog are  
CC involved in the maturation of T lymphocytes. The specification describes  
CC a method for modulating immune function, by administration of a hedgehog  
CC or patched (ptc) polypeptide, agonists or antagonists. The method is used  
CC to treat disorders affecting the regulation of lymphocytes, particularly  
CC maturation and/or activation of T lymphocytes. It is used to treat  
CC bacterial or viral infection, diabetes, nutritional deficiencies, graft  
CC rejection or other hyperacute response such as kidney, heart, lung, bone  
CC marrow spleen skin or cornea transplant or autoimmune disorders such as  
CC multiple sclerosis, psoriasis or atopic dermatitis. The method is used to  
CC treat inflammatory, proliferative and hyperproliferative diseases, as  
CC well as cutaneous manifestations of immunological disorders such as  
CC eczematous dermatitis, urticaria, vasculitis and scleroderma  
XX  
SQ Sequence 475 AA;  
Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLARCLLLVLSLLVCSGLACGPGFGRKRRPKLTPLAYKQIPNVAEKTLCASG 60  
Db 1 MLLARCLLLVLSLLVCSGLACGPGFGRKRRPKLTPLAYKQIPNVAEKTLCASG 60  
QY 61 RYEGKISNSRFRKELTPNYPDIIFKDEVTGADRLMTQCKLNALAISVNNQPGV 120  
Db 61 RYEGKISNSRFRKELTPNYPDIIFKDEVTGADRLMTQCKLNALAISVNNQPGV 120  
QY 121 KLRTVEGHWDEGHSESLHYEGRAVDITTSDDRSKYGMRLAVRAGFDWVYESKAH 180  
Db 121 KLRTVEGHWDEGHSESLHYEGRAVDITTSDDRSKYGMRLAVRAGFDWVYESKAH 180  
QY 181 IHCSVKAENSVAAKSGGCPGSAVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAKSGGCPGSAVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKYVYIETREPERILLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKYVYIETREPERILLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPQRYVYVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPQRYVYVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCVAVIEHWAHRAFPFLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
Db 361 ASCVAVIEHWAHRAFPFLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGTHWSQLLYQIGTWLLDSEALHPLGNVKSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGTHWSQLLYQIGTWLLDSEALHPLGNVKSXSRGAGGAREGA 475

RESULT 11  
AAB60265  
ID AAB60265 standard; protein; 475 AA.  
XX

AC AAB60265;  
XX 30-MAR-2001 (First entry)  
XX Human Sonic hedgehog (Shh) protein, SEQ ID NO:15.  
DE Hedgehog protein; polymer conjugate; polyalkene glycol group;  
XX bioavailability; formulation; neurological disorder;  
KW inflammatory disorder; autoimmune disorder; cancer;  
KW neurodegenerative disorder; Parkinson's disease; Huntington's disease;  
KW Alzheimer's disease; neurological injury; stroke; multiple sclerosis;  
KW malignant glioma; medulloblastoma; neuroectodermal tumour.  
XX  
OS Homo sapiens.  
XX WO200073337-A1.  
PN 07-DEC-2000.  
PD 26-MAY-2000; 2000WO-US014741.  
XX 01-JUN-1999; 99US-0137011P.  
PR 13-AUG-1999; 99US-0149016P.  
XX (BIOJ) BIOGEN INC.  
PA  
XX Pepinsky RB, Taylor F, Garber E;  
PI WPI; 2001-049927/06.  
XX N-PSDB; AAP27018.  
DR  
XX Modified hedgehog protein, useful in the treatment of Parkinson's disease  
PT and Huntington's chorea, comprises a polymer containing a polyalkylene  
PT glycol group linked to any residue other than the N-terminal and lysine  
PT residues.  
XX  
PS Disclosure; Page 138-140; 157pp; English.  
XX The invention relates to novel polymer conjugates of hedgehog proteins  
CC which have increased bioavailability. The hedgehog proteins are  
CC conjugated to a non-naturally-occurring polymer comprising a polyalkylene  
CC glycol group, with the proviso that the polymer is not conjugated to the  
CC N-terminus, or to lysine residues of the hedgehog protein. The hedgehog  
CC protein used in the conjugate may be a wild-type or mutant Sonic hedgehog  
CC (Shh), Indian hedgehog (Ihh) or Desert hedgehog (Dhh) protein, or may be  
CC a hedgehog fusion protein. The invention also relates to methods of  
CC defining and mapping functionally important regions of a protein by  
CC modifying accessible amino acid side chains, and determining the effect  
CC the position and/or type of modification have on the activity of the  
CC protein. The hedgehog polymer conjugates may be used in the management of  
CC various medical conditions including various neurological disorders,  
CC inflammatory and autoimmune diseases, and cancers. In particular, they  
CC may be used to prevent preventing or ameliorate neurodegenerative  
CC disorders (e.g., Parkinson's disease, Huntington's disease, Alzheimer's  
CC disease); age-associated neurological disease; neurological injury and  
CC trauma; immunological diseases of the nervous system (e.g., multiple  
CC sclerosis); stroke; and malignant gliomas, medulloblastomas and  
CC neuroectodermal tumours. The modifications made to the hedgehog protein  
CC may result in increased half-life, altered tissue distribution (such as  
CC an improved ability to stay in the vasculature for longer periods of  
CC time), increased stability in solution, protection from proteolytic  
CC degradation, or reduced immunogenicity. In particular, the ability to  
CC remain in the vasculature for prolonged periods may allow a hedgehog  
CC protein of the invention to cross the blood-brain barrier, and an  
CC increased thermal stability would be an advantage when formulating the  
CC hedgehog protein in powder form. The present sequence represents a member  
CC of the hedgehog family of proteins  
XX  
SQ Sequence 475 AA;  
Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTIGASG 60  
 QY 61 RYEGKISRNSERFKELTPNNPDIIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 120  
 Db 61 RYEGKISRNSERFKELTPNNPDIIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 120  
 QY 121 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYESKAH 180  
 Db 121 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYESKAH 180  
 QY 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 QY 301 PRALFASVRPGQRYVYVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASVRPGQRYVYVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
 QY 361 ASCYAVIEHSHWAHRAPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGGAREGA 475

RESULT 12  
 AA04687  
 ID AA04687 standard; protein; 475 AA.  
 XX  
 AC AA04687;  
 DT 04-SEP-2001 (first entry)  
 DE Human sonic hedgehog (Shh) protein.  
 KW Human; hedgehog protein; nootropic; neuroprotective; anticonvulsant;  
 KW cytosolic; therapy; Alzheimer's disease; Parkinson's disease; injury;  
 KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;  
 KW nervous system aging; neurodegenerative disease; immunological disease;  
 KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;  
 KW extracellular signalling protein.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 24..137  
 FT Region /note= "N-terminal fragment"  
 FT Misc-difference 463  
 FT /label= Unknown  
 FT /note= "Encoded by NNN"  
 XX  
 PN W0200134654-A1.  
 XX  
 XX 17-MAY-2001.  
 XX  
 XX 02-NOV-2000; 2000WO-US030405.  
 XX  
 XX 05-NOV-1999; 99US-0164025P.  
 XX  
 XX (BIOJ ) BIOGEN INC.  
 XX  
 XX Strauch K;  
 PI  
 XX  
 XX WPI; 2001-329075/34.  
 DR

DR N-PSDB; AAD09034.  
 XX Novel isolated hedgehog fusion polypeptide useful for treating  
 PT neurological conditions such as Alzheimer's disease, Parkinson's disease,  
 PT Huntington's chorea, amyotrophic lateral sclerosis, and multiple  
 XX sclerosis.  
 XX Disclosure; Page 117-118; 178pp; English.  
 CC The present invention relates to hedgehog fusion proteins. Hedgehog  
 CC proteins are a family of extracellular signalling proteins that regulate  
 CC various aspects of embryonic development both in vertebrates and in  
 CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or  
 CC treatment of any condition or disease state for which a hedgehog or  
 CC patched protein constituent is efficacious and in the diagnosis of  
 CC constituents or conditions of disease states in biological system or  
 CC specimens and for diagnostic purposes in non-physiological systems.  
 CC Hedgehog fusion protein is useful for treating neurological conditions  
 CC due to injury, aging of nervous system, including Alzheimer's disease,  
 CC chronic neurodegenerative diseases of the nervous system, including  
 CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis  
 CC and chronic immunological diseases of nervous system including multiple  
 CC sclerosis and malignant gliomas, medulloblastomas, neuroectodermal  
 CC tumours and to specifically target medical therapies against cancer and  
 CC tumours which express the receptor for the protein. The present sequence  
 CC is human sonic hedgehog (Shh) protein  
 XX  
 SQ Sequence 475 AA;  
 Query Match 99.9%; Score 2467; DB 4; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTIGASG 60  
 QY 61 RYEGKISRNSERFKELTPNNPDIIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 120  
 Db 61 RYEGKISRNSERFKELTPNNPDIIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 120  
 QY 121 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYESKAH 180  
 Db 121 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYESKAH 180  
 QY 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 QY 301 PRALFASVRPGQRYVYVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASVRPGQRYVYVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
 QY 361 ASCYAVIEHSHWAHRAPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGGAREGA 475

RESULT 13  
 AA05377  
 ID AA05377 standard; protein; 475 AA.  
 XX  
 AC AA05377;  
 XX  
 DT 12-SEP-2001 (first entry)



CC The invention provides a method for regulating formation and/or  
 CC maintenance of adipocyte tissue that comprises contacting pre adipocyte  
 CC or adipocyte cells with a hedgehog polypeptide or ptc therapeutic. The  
 CC method is used for regulating the growth state of an adipocyte stem/  
 CC progenitor cell, and treating or preventing disorders of, or surgical or  
 CC cosmetic repair of, adipocyte tissues, e.g. for treating or preventing  
 CC hyperplastic or neoplastic conditions affecting adipocyte tissue, such as  
 CC soft tissue tumors, especially adipose cell tumors, e.g. lipomas,  
 CC fibrolipomas, lipoblastomas, lipomatosis, hibernomas, hemangiomas and/or  
 CC liposarcomas. Hedgehog polypeptides can be used in combination with other  
 CC therapeutic agents. The present sequence represents a human sonic  
 CC hedgehog (Shh) polypeptide  
 XX  
 XX Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 4; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSVLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVSVLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60

QY 61 RYEGKISRNSRFKELTNNYNDIIFKDEENTGADRLMTORCKDKNALAISVMNQPGV 120  
 Db 61 RYEGKISRNSRFKELTNNYNDIIFKDEENTGADRLMTORCKDKNALAISVMNQPGV 120

QY 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180  
 Db 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180

QY 181 IHCSVKAENSVAAKSGCGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAKSGCGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGPEASSGGSPGGALG 300  
 Db 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGPEASSGGSPGGALG 300

QY 301 PRALFASRVRCQRYVVAERDGRRLPAAVHVTI SEEAAGAYAPLTAGTITLINRVL 360  
 Db 301 PRALFASRVRCQRYVVAERDGRRLPAAVHVTI SEEAAGAYAPLTAGTITLINRVL 360

QY 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRGGRVALTPGA 420  
 Db 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRGGRVALTPGA 420

QY 421 ADAPGAGATAGIHWYSQLLYQITGTLWLDSEALHPLGMVAKSXSRGAGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQITGTLWLDSEALHPLGMVAKSXSRGAGGAREGA 475

## RESULT 15

ID AAB31222 standard; protein; 475 AA.  
 XX AAB31222;  
 AC AAB31222;  
 XX  
 DT 20-APR-2001 (first entry)  
 XX

XX Amino acid sequence of human sonic hedgehog protein (Shh).

XX Hedgehog related-protein; sonic hedgehog protein; Shh; ischemia; stroke;  
 KW desert hedgehog protein; Dhh; indian hedgehog protein; inh; neuron;  
 KW neurological condition; nervous system injury; tumor-induced injury;  
 KW aging; Alzheimer's disease; chronic neurodegenerative disease;  
 KW Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;  
 KW spinocerebellar degeneration; chronic immunological disease;  
 KW multiple sclerosis.

XX Homo sapiens.

OS  
 XX Key Location/Qualifiers  
 FH

FT Misc-difference 463 /note= "unspecified amino acid encoded by NNN"  
 FT US6165747-A.  
 PN 26-DEC-2000.  
 XX 05-JUN-1995; 95US-00460900.  
 XX 30-DEC-1993; 93US-00176427.  
 PR 14-DEC-1994; 94US-00356060.  
 PR 04-MAY-1995; 95US-00435033.  
 XX (HARD ) HARVARD COLLEGE.  
 PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 XX Ingham PW, McMahon AP, Tabin CJ, Marti-Gorostiza E, Bumcrot DA;  
 XX WPI: 2001-079847/09.  
 DR N-PSDB; AAC87079.

XX Polynucleotides encoding hedgehog proteins, useful for treating diseases  
 of nervous system such as Alzheimer's disease, Parkinson's disease,  
 Huntington's chorea, amyotrophic lateral sclerosis, multiple sclerosis.  
 PS Claim 10; Col 147-150; 119pp; English.

XX The present sequence represents a hedgehog related-protein. The  
 CC specification describes a sonic hedgehog protein (Shh), a desert hedgehog  
 CC protein (Dhh), and an indian hedgehog protein (Ihh). The hedgehog  
 CC polynucleotides are useful in diagnostic, in antisense therapy and in  
 CC therapeutic assays for detecting and treating disorders involving, e.g.,  
 CC aberrant expression of hedgehog and hedgehog homologues. Hedgehog  
 CC polypeptides are useful therapeutically to enhance survival of neurons  
 CC and other neuron cells and in treating neurological conditions deriving  
 CC from acute, subacute, or chronic injury to the nervous system, including  
 CC traumatic injury, chemical injury, vascular injury and deficits (such as the  
 CC ischemia resulting from stroke), together with infectious/inflammatory  
 CC and induced-induced injury, aging of the nervous system including  
 CC Alzheimer's disease, chronic neurodegenerative diseases of the nervous  
 CC system, including Parkinson's disease, Huntington's chorea, amyotrophic  
 CC lateral sclerosis, spinocerebellar degenerations, and chronic  
 CC immunological diseases of the nervous system or affecting the nervous  
 CC system, including multiple sclerosis

XX Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 4; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSVLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVSVLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60

QY 61 RYEGKISRNSRFKELTNNYNDIIFKDEENTGADRLMTORCKDKNALAISVMNQPGV 120  
 Db 61 RYEGKISRNSRFKELTNNYNDIIFKDEENTGADRLMTORCKDKNALAISVMNQPGV 120

QY 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180  
 Db 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYVESKAH 180

QY 181 IHCSVKAENSVAAKSGCGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAKSGCGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGPEASSGGSPGGALG 300  
 Db 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGPEASSGGSPGGALG 300

QY 301 PRALFASRVRCQRYVVAERDGRRLPAAVHVTI SEEAAGAYAPLTAGTITLINRVL 360  
 Db 301 PRALFASRVRCQRYVVAERDGRRLPAAVHVTI SEEAAGAYAPLTAGTITLINRVL 360

Db 301 PRALFASRVQRVYVVAERDGRRLPAAVSHVTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTDRGGSGGGGRVALTAPGA 420  
Db 361 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTDRGGSGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 16  
ABB79137  
ID ABB79137 standard; protein; 475 AA.  
XX AC ABB79137;  
XX DT 06-AUG-2002 (first entry)  
XX DE Human sonic hedgehog (Shh) protein SEQ ID NO:13.  
XX KW Sonic hedgehog; Shh; desert hedgehog; Dh; Indian hedgehog; Ihh;  
KW antiparkinsonian; antiarrhythmic; neuroprotective; anticonvulsant;  
KW cytosolic; nootropic; spermatogenesis; peripheral nervous system;  
KW central nervous system; Alzheimer's disease; Parkinson's disease;  
KW Huntington's disease; arrhythmia; nerve degeneration; multiple sclerosis;  
KW immunological disorder; neoplastic; hyperplastic.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT Misc-difference 463  
FT /label= unknown  
FT /note= "encoded by NNN"  
XX US6384192-B1.  
XX 07-MAY-2002.  
XX 20-OCT-1997; 97US-00957874.  
XX 30-DEC-1993; 93US-00176427.  
XX 14-DEC-1994; 94US-00356050.  
XX 04-MAY-1995; 95US-00435093.  
XX 05-JUN-1995; 95US-00462386.  
XX (HARD ) HARVARD COLLEGE.  
XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX Ingham PW, McMahon AP, Tabin CJ;  
XX WPI; 2002-442817/47.  
XX N-PSDB; ABN87549.  
XX New vertebrate hedgehog-related proteins, useful e.g. for promoting  
XX differentiation, survival and proliferation of cells, e.g. for treating  
XX neurodegeneration.  
XX Claim 2; Col 149-152; 116pp; English.

XX The present invention describes an isolated and/or recombinant  
XX polypeptide (I) comprising a hedgehog (hh) amino acid (aa) sequence  
XX encoded by a nucleic acid (II) that hybridizes under stringent conditions  
XX to 1 of 6 sequences (see ABN87544, and ABN87546 to ABN87550). (I) binds  
XX to a natural patched receptor. Specifically claimed example of (I) are  
XX given in ABB79132 and ABB79134 to ABB79138. (I) has antiparkinsonian,  
XX neurotrophic, neuroprotective, anticonvulsant, antiarrhythmic and cytosolic  
XX activities. (I) induces the expression of the BMP-2 and -4 genes, and of  
XX the Hoxd gene. (I) can be used: (i) to promote differentiation of  
XX neuronal cells and survival of the differentiated cells, specifically  
XX dopaminergic or motor neurons, proliferation of chondrocytes, and  
XX proliferation, differentiation and/or survival of mesodermal or  
XX ectodermal cells, either in cell cultures (particularly for preparation

CC of transplants) or therapeutically; (ii) for detecting loss of response,  
CC in tissues or cells, to hh proteins; (iii) in drug screening (to identify  
CC antagonists, useful e.g. for inhibition of spermatogenesis); and (iv)  
CC for isolation of cognate receptors. (I) may be used therapeutically to  
CC treat e.g. injuries/defects in the central or peripheral nervous systems,  
CC including Alzheimer's, Parkinson's and Huntington's diseases, or  
CC arrhythmias caused by nerve degeneration; immunological disorders of the  
CC nervous system, e.g. multiple sclerosis; neoplastic and hyperplastic  
CC alterations in the central nervous system, also to promote attachment of  
CC prospheres. The present sequence represents the human sonic hedgehog  
XX (Shh) protein given in the present invention  
XX SQ Sequence 475 AA;  
XX Query Match 99.9%; Score 2467; DB 5; Length 475;  
XX Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
XX Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLARCLLLVSVLLVCSGLACGPGRGFGKRRHPKLLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVSVLLVCSGLACGPGRGFGKRRHPKLLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDBENTGADRLMTQRCCKDKNALAI SVMNQWPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDBENTGADRLMTQRCCKDKNALAI SVMNQWPGV 120  
QY 121 KLRVTEGWDDEGHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 121 KLRVTEGWDDEGHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
QY 181 IHCSVKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGLLYSDFLT 240  
Db 181 IHCSVKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGLLYSDFLT 240  
QY 241 FLDRDDGAKKVFVVIETREPRERLLITAAHLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFVVIETREPRERLLITAAHLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASRVQRVYVVAERDGRRLPAAVSHVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVQRVYVVAERDGRRLPAAVSHVTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTDRGGSGGGGRVALTAPGA 420  
Db 361 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTDRGGSGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 17  
AAE14297  
ID AAE14297 standard; protein; 475 AA.  
XX AC AAE14297;  
XX DT 07-MAR-2002 (first entry)  
XX DE Human Sonic hedgehog (Shh) protein.  
XX KW Neuronal degeneration; hedgehog therapeutic; neurotrophic factor;  
KW therapy; amelioration; Alzheimer's disease; Parkinson's disease;  
KW Huntington's chorea; amyotrophic lateral sclerosis; schizophrenia;  
KW neurological disorder; hypoglycaemia; senile dementia; depression;  
KW Korsakoff's disease; acquired immune deficiency syndrome; hypoxia;  
KW AIDS dementia; epileptic seizure; Tourette's syndrome; ischaemia;  
KW suffocation; neurotoxic injury; cerebrovascular accident; anoxia; trauma;  
KW myocardial infarct; drowning; perinatal asphyxia; surgery;  
KW drug addiction; stroke; human; Sonic hedgehog; Shh protein; ALS.  
XX OS Homo sapiens.

PH Key Location/Qualifiers  
PT Region 24..197  
PT /notes="N-terminal fragment; This region is specifically  
PT referred in claim 3"  
PT Misc-difference 463  
PT /note= "Encoded by NNN"  
XX  
XX WO200182946-A2.  
XX  
XX 08-NOV-2001.  
XX  
XX 27-APR-2001; 2001WO-US013854.  
XX  
XX 28-APR-2000; 2000US-0200765P.  
XX  
XX (CURI-) CURIS INC.  
XX  
XX Reilly JO;  
XX  
XX WPI; 2002-034488/04.  
XX  
XX N-PSDB; AAD23804.  
XX  
XX  
XX Ameliorating neuronal degeneration such as Huntington's Chorea,  
PT Alzheimer's or Parkinson's disease in a subject, by administering to the  
PT subject a composition comprising hedgehog therapeutic and neurotrophic  
PT factor.  
XX  
XX Claim 3; Page 127-128; 139pp; English.  
XX  
XX The invention relates to a method of ameliorating neuronal degeneration  
CC in a subject. The method involves administering to the subject a  
CC composition comprising hedgehog therapeutic and a neurotrophic factor.  
CC The method is useful for the treatment, amelioration or prevention of  
CC Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis  
CC (ALS) and Huntington's chorea. It is useful for promoting the survival  
CC and/or differentiation of cholinergic neurons. The method is also useful  
CC for treating or preventing neurological disorders such as hypoglycaemia,  
CC senile dementia, Korsakoff's disease, schizophrenia, acquired immune  
CC deficiency syndrome (AIDS) dementia, depression, epileptic seizures,  
CC Tourette's syndrome, drug addiction, for reducing neurotoxic injury  
CC associated with conditions of hypoxia, anoxia or ischaemia which  
CC typically follows stroke, cerebrovascular accident, brain or spinal cord  
CC trauma, myocardial infarct, physical trauma, drowning, suffocation,  
CC perinatal asphyxia, hypoglycaemic events, and for treating or preventing  
CC the adverse neurological consequences of surgery. The present sequence is  
CC human Sonic hedgehog (Shh) protein  
XX  
XX Sequence 475 AA;  
SQ

Query Match 99.9%; Score 2467; DB 5; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.le-217;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLARCLLVLSVLLVCSGLACGPGRGFKRRHPKKLTPLAYKQIPNVAEKTILGASG 60  
DB 1 MLLARCLLVLSVLLVCSGLACGPGRGFKRRHPKKLTPLAYKQIPNVAEKTILGASG 60  
QY 61 RYEGKISRNSRPFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAISSVNNQWPGV 120  
DB 61 RYEGKISRNSRPFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAISSVNNQWPGV 120  
QY 121 KLRVTEGWDECHHSSESLHVEGRAVDITTSDDRSKYGMARLAVBAGFDWVYVESKAH 180  
DB 121 KLRVTEGWDECHHSSESLHVEGRAVDITTSDDRSKYGMARLAVBAGFDWVYVESKAH 180  
QY 181 IHCSVKAENSVAAGSGGCPGSAVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAGSGGCPGSAVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKYFYVITETREPRERILLTAHLLFVAPENDSATGPEASSGSGPPSGGALG 300  
DB 241 FLDRDDGAKKYFYVITETREPRERILLTAHLLFVAPENDSATGPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPRGQRYVYVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASRVPRGQRYVYVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDGGSDGGGDRGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDGGSDGGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGNMAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGNMAVKSSXSRGAGGAREGA 475  
RESULT 18  
AAO20923  
ID AAO20923 standard; protein; 475 AA.  
XX  
XX AAO20923;  
XX  
XX 08-JUL-2002 (first entry)  
XX  
XX Human Shh protein sequence SEQ ID No 15.  
XX  
XX Cytostatic; ophthalmological; antirheumatic; antiarthritic; osteopathic;  
KW gastrointestinal general; antiinflammatory; antiulcer; antipsoriatic;  
KW vulnerary; vasotropic; antidiabetic; cerebroprotective; immunomodulator;  
KW muscular active general; hypotensive; antilipaeamic; hedgehog protein;  
KW contraceptive; antiinfertility; angiogenesis; placentation; retinopathy;  
KW malignant tumour; macular degeneration; non-malignant tumour; keloid;  
KW rheumatoid arthritis; osteoarthritis; neovascular glaucoma; psoriasis;  
KW Crohn's disease; ulcerative colitis; tissue repair; diabetic retinopathy;  
KW ischaemia; inflammation; peripheral; central nervous system; cachexia;  
KW vascular disease; high blood pressure; cholesterol; ovulation; cancer;  
KW gynaecological; menstruation; endometrial lining formation.  
XX  
XX Homo sapiens.  
XX  
XX WO200198344-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 18-JUN-2001; 2001WO-US019435.  
XX  
XX 16-JUN-2000; 2000US-0211919P.  
XX  
XX (BIOJ ) BIOGEN INC.  
XX  
XX Ling LE, Sanicola-Nadel M;  
XX  
XX WPI; 2002-291693/33.  
XX  
XX N-PSDB; AAK99685.  
XX  
XX Use of hedgehog polypeptide, its agonists or antagonists for modulating  
PT angiogenesis in the treatment of e.g. malignant tumors, retinopathy,  
PT macular degeneration, non-malignant tumors, rheumatoid arthritis and  
PT osteoarthritis.  
XX  
XX Claim 4; Page 258-259; 269pp; English.  
XX  
XX The invention relates to new methods for modulating (promoting or  
CC inhibiting) angiogenesis in a subject animal using hedgehog polypeptides  
CC or its modulators. The hedgehog polypeptide or its modulators are useful  
CC for modulating angiogenesis in a subject. The antagonists are useful in  
CC the treatment of angiogenesis related disorders such as malignant  
CC tumours, retinopathy, macular degeneration, non-malignant tumours,  
CC rheumatoid arthritis, osteoarthritis, neovascular glaucoma, keloids, of the  
CC Crohn's disease, ulcerative colitis and psoriasis. The antagonists of the  
CC invention are also useful for the treatment of tissue repair, ischaemia,  
CC diabetic retinopathy, inflammation, peripheral or central nervous system  
CC vascular disease, cachexia and high blood pressure and cholesterol levels  
CC ; for the modulation of ovulation, menstruation, placentation and  
CC endometrial lining formation and maintenance; and for the diagnosis of  
CC e.g. cancer. This sequence is a hedgehog protein of the invention

XX SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 5; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60

Qy 61 RYEGKISRNSRPFKELTNYNPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Db 61 RYEGKISRNSRPFKELTNYNPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120

Qy 121 KLRVTEGDEDDGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYESKAH 180  
 Db 121 KLRVTEGDEDDGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYESKAH 180

Qy 181 IHCSVKAENSVAAKSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAKSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300

Qy 301 PRALFASVRPQGVYVVAERDGRRLLPAAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASVRPQGVYVVAERDGRRLLPAAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360

Qy 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAPGA 420

Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGAREGA 475

RESULT 19  
 AAU99483  
 ID AAU99483 standard; protein; 475 AA.  
 AC AAU99483;  
 XX 07-OCT-2002 (first entry)  
 DT Human sonic hedgehog (Shh) protein.  
 DE Basal cell carcinoma; BCC; non-carcinoma skin sample; dermatitis; epidermis;  
 KW culture medium; hedgehog signalling pathway; growth; development;  
 KW abnormality; cancer; cytostatic; human; Shh; sonic hedgehog.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference /label= Unknown  
 FT  
 FT  
 XX WO200244344-A2.  
 XX  
 XX 06-JUN-2002.  
 XX  
 XX 28-NOV-2001; 2001WO-US044457.  
 XX  
 XX 28-NOV-2000; 2000US-0253660P.  
 XX (CURTI-) CURTI INC.  
 PA Berg AP, Pepicelli C, Wang F;  
 PI WPI; 2002-557537/59.  
 XX  
 XX

DR N-PSDB; ABK88636.  
 XX Preparing basal cell carcinoma culture useful for identifying BCC  
 PT therapeutics by contacting non-carcinoma skin sample having dermis and  
 PT epidermis, with culture medium having hedgehog agonist and/or hedgehog  
 XX protein.  
 XX Disclosure; Page 107-108; 117pp; English.  
 XX  
 XX The present invention relates to methods for preparing basal cell  
 CC carcinoma (BCC) cultures. The method involves obtaining a non-carcinoma  
 CC skin sample comprising dermis and epidermis and placing it in contact  
 CC with culture medium comprising a hedgehog agonist and/or hedgehog  
 CC protein. The method optionally involves either obtaining a BCC sample and  
 CC placing it in close proximity to a dermis sample, or obtaining a BCC and  
 CC culturing the BCC sample at a liquid/air interface. The BCC cultures and  
 CC the methods of the invention are useful for identifying a BCC therapeutic  
 CC (e.g. a hedgehog agonist or antagonist) and allowing more extensive in  
 CC vitro studies of BCCs. Since BCC characteristics depend heavily upon  
 CC hedgehog pathway activity, BCC cultures may be used to identify compounds  
 CC that regulate the hedgehog signalling pathway. Compounds that affect BCC  
 CC growth and development may also be useful in treating a wide range of  
 CC diseases, disorders and abnormalities. The cultures may be used to  
 CC examine essentially any aspect of BCC biology in vitro, thus leading to  
 CC the discovery of valuable new therapeutics for treatment of this most  
 CC common form of cancer. The present sequence represents human sonic  
 CC hedgehog (Shh) protein  
 XX  
 XX Sequence 475 AA;  
 SQ

Query Match 99.9%; Score 2467; DB 5; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60

Qy 61 RYEGKISRNSRPFKELTNYNPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Db 61 RYEGKISRNSRPFKELTNYNPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQPGV 120

Qy 121 KLRVTEGDEDDGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYESKAH 180  
 Db 121 KLRVTEGDEDDGHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYESKAH 180

Qy 181 IHCSVKAENSVAAKSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAKSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300

Qy 301 PRALFASVRPQGVYVVAERDGRRLLPAAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASVRPQGVYVVAERDGRRLLPAAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360

Qy 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAPGA 420

Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVKSXSXRGAGGAREGA 475

RESULT 20  
 ADA26256  
 ID ADA26256 standard; protein; 475 AA.  
 XX  
 XX ADA26256;  
 XX

20-NOV-2003 (first entry)  
Human Sonic hedgehog (Shh) polypeptide.  
Human; Sonic hedgehog; Shh; neuronal cell; skeletogenesis;  
chondrogenesis; osteogenesis; degenerative disorder; nervous system;  
neuronal cell death; neural cell; neuromuscular disorder;  
autonomic disorder; central nervous system disorder; anoxia; ischaemia;  
peripheral nervous system disorder; tachycardia;  
atrial cardiac arrhythmia; striated heart; stem cell development;  
digestive tract; liver; multiple sclerosis.  
Homo sapiens.  
Key Location/Qualifiers  
Misc-difference 463  
/label= unknown  
US2003054437-A1.  
20-MAR-2003.  
20-OCT-1997; 97US-00954771.  
30-DEC-1993; 93US-00176427.  
14-DEC-1994; 94US-00356060.  
04-MAY-1995; 95US-00435093.  
05-JUN-1995; 95US-00462386.  
(INGH/) INGHAM P W.  
(MCMH/) MCMAHON A P.  
(TAB1/) TABIN C J.  
Ingham PW, McMahon AP, Tabin CJ;  
WPI; 2003-555377/52.  
N-PSDB; ADA26295.  
Modulating growth, differentiation or survival of a cell, useful for  
treating a degenerative disorder of the nervous system characterized by  
neuronal cell death, comprises contacting the cell with a hedgehog  
polypeptide.  
Claim 5; Page 81-82; 121pp; English.  
The invention relates to a method for modulating growth, differentiation  
or survival of a cell, comprising contacting the cell with a hedgehog  
polypeptide. The invention also relates to methods for inducing a cell to  
differentiate to a neuronal cell phenotype comprising contacting the cell  
with a hedgehog polypeptide, modulating skeletogenesis by contacting a  
target tissue of a hedgehog polypeptide to cause chondrogenesis and/or  
osteogenesis in the target tissue and treating a degenerative disorder of  
the nervous system characterised by neuronal cell death, comprising  
administering a hedgehog polypeptide causing prolonged survival of neural  
cells in the patient, relative to the absence of hedgehog treatment. The  
hedgehog polypeptides are useful for treating a degenerative disorder of  
the nervous system characterised by neuronal cell death, including  
neuromuscular, autonomic or central nervous system disorders,  
specifically Alzheimer's disease, Parkinson's disease, amyotrophic  
lateral sclerosis, Pick's disease, Huntington's disease, multiple  
sclerosis, neuronal damage resulting from anoxia, ischaemia or trauma and  
neuronal degeneration associated with a natural aging process. The  
polypeptides may also be used for treating peripheral nervous system  
disorders including disorders affecting innervation of smooth muscle and  
endocrine tissue, such as tachycardia or atrial cardiac arrhythmias which  
may arise from a degenerative condition whereby the nerves innervate the  
striated muscle of the heart, in nerve prostheses for repairing central  
and peripheral nerve damage, for treating neoplastic or hyperplastic  
transformations and in controlling the development of stem cells  
responsible for the formation of the digestive tract, liver and other  
organs. This sequence represents the human Sonic hedgehog (Shh)  
polypeptide.

SQ Sequence 475 AA;  
Query Match 99.98; Score 2467; DB 6; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-217; Indels 0; Gaps 0;  
Matches 475; Conservative 0; Mismatches 0;  
QY 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLLTPLAYKQIPNVAEKTILGASG 60  
DB 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLLTPLAYKQIPNVAEKTILGASG 60  
QY 61 RYEGKISRNSERFKELTPNYPDIIIFKDEENTGADRLMTORCKDKLNALAISSVNNQPGV 120  
DB 61 RYEGKISRNSERFKELTPNYPDIIIFKDEENTGADRLMTORCKDKLNALAISSVNNQPGV 120  
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
DB 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
QY 181 IHCSVKAENSVAAGSGCGFCPSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAGSGCGFCPSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLLTAAHLLFVAPHNDSATGPEPEASSGSGPSSGGAIG 300  
DB 241 FLDRDDGAKKVFYVETREPRERLLLTAAHLLFVAPHNDSATGPEPEASSGSGPSSGGAIG 300  
QY 301 PRALFASRVRFQGVVVAERDGRRLPPAAVHVSITLSEENAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASRVRFQGVVVAERDGRRLPPAAVHVSITLSEENAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGGSGGGGGRGGRVALTAPGA 420  
DB 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGGSGGGGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYOIGTWLLDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWSQLLYOIGTWLLDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
RESULT 21  
ADD25260  
ID ADD25260 standard; protein; 475 AA.  
XX  
XX AC ADD25260;  
XX DT 15-JAN-2004 (first entry)  
XX DE Human Sonic hedgehog (Shh) polypeptide.  
XX KW human; Sonic hedgehog; SHH; patched receptor; spermatogenesis inhibition;  
KW ovary function inhibition; embryogenesis;  
KW differential tissue maintenance.  
XX OS Homo sapiens.  
XX PN US6576237-B1.  
XX PD 10-JUN-2003.  
XX PF 16-AUG-2000; 2000US-00639695.  
XX PR 30-DEC-1993; 93US-00176427.  
XX PR 14-DEC-1994; 94US-00356060.  
XX PR 04-MAY-1995; 95US-00435093.  
XX PR 05-JUN-1995; 95US-00460900.  
XX PA (HARD ) HARVARD COLLEGE.  
XX PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX PI Ingham PW, McMahon AP, Tabin CJ, Bumerot DA, Marti-Gorostiza E;  
XX DR WPI; 2003-799823/75.  
XX DR N-PSDB; ADD25315.

XX Novel isolated antibody which is immunoreactive with a vertebrate  
 PT hedgehog protein sequence that binds with patched receptor, useful for  
 PT blocking action of naturally occurring hedgehog protein, and for  
 PT inhibiting spermatogenesis.  
 XX  
 PS Claim 6; SEQ ID NO 13; 120pp; English.  
 XX  
 CC The invention relates to an isolated antibody (I) which is immunoreactive  
 CC with a hedgehog polypeptide (II) that binds to a patched receptor, where  
 CC (II) is encoded by nucleic acid which hybridize to a fully defined  
 CC vertebrate hedgehog (hh) protein. (I) is useful as a hedgehog antagonist  
 CC by blocking action of naturally occurring hedgehog protein, and therefore  
 CC for inhibiting spermatogenesis. (I) is also useful for inhibiting normal  
 CC ovarian function. (I) is useful for blocking the action of one or more  
 CC hedgehog proteins and allows the study of the role of these proteins  
 CC e.g., embryogenesis and/or maintenance of differential tissue. (I) is  
 CC also useful in immunohistochemical staining of tissue samples in order to  
 CC evaluate the abundance and pattern of expression of the hedgehog  
 CC polypeptides. (I) is also useful diagnostically in immunoprecipitation  
 CC and immunoblotting to detect and evaluate hedgehog protein levels as a  
 CC part of clinical testing procedure. The present sequence represents the  
 CC amino acid sequence of a hedgehog polypeptide.  
 XX  
 SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 7; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGGFGKRRHPKLTPLAYKQIPNVAEKLTCASG 60  
 DB 1 MLLARCLLLVSSLLVCSGLACGPGGFGKRRHPKLTPLAYKQIPNVAEKLTCASG 60  
 QY 61 RYEGKISNSRERFELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQWPGV 120  
 DB 61 RYEGKISNSRERFELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQWPGV 120  
 QY 121 KLRTVEGWDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
 DB 121 KLRTVEGWDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
 QY 181 IHCSVKAENSVAAKSGGCGFPASATVHLPGQGTCLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
 DB 181 IHCSVKAENSVAAKSGGCGFPASATVHLPGQGTCLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
 QY 241 FLDRDDGAKVPYVETREPRERILLTAHLLFVAPHNDSATGPEASSGSGPSPGGALG 300  
 DB 241 FLDRDDGAKVPYVETREPRERILLTAHLLFVAPHNDSATGPEASSGSGPSPGGALG 300  
 QY 301 PRALFASVRPQQRVYVVAERDGRLLPAAVHVSVTLSSEAGAYAPLTAQTILINRVL 360  
 DB 301 PRALFASVRPQQRVYVVAERDGRLLPAAVHVSVTLSSEAGAYAPLTAQTILINRVL 360  
 QY 361 ASCYAVIEHSHWAHFAFPFRLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAPGA 420  
 DB 361 ASCYAVIEHSHWAHFAFPFRLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAPGA 420  
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKVXSXRGAGGGAREGA 475  
 DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKVXSXRGAGGGAREGA 475

RESULT 22  
 ID ABW00868  
 XX ABW00868 standard; protein; 475 AA.  
 AC ABW00868;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX Human sonic hedgehog protein.  
 DE  
 XX

KW Human; cell differentiation; Desert hedgehog; Dhh; Sonic hedgehog; shh;  
 KW Indian hedgehog; Ihh; skeletogenesis; degenerative disorder; ischaemia;  
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
 KW Huntington's disease; multiple sclerosis; Pick's disease; aging process;  
 KW trauma; anoxia; antisense gene therapy; neuroprotective; anticonvulsant;  
 KW neurotropic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 463 /note= "Encoded by NNN"  
 XX  
 XX US2003186357-A1.  
 XX 02-OCT-2003.  
 XX 05-JUN-1995; 95US-00462386.  
 XX 30-DEC-1993; 93US-00176427.  
 XX 14-DEC-1994; 94US-00356060.  
 XX 04-MAY-1995; 95US-00435093.  
 XX (INGH/) INGHAM P W.  
 XX (MCMA/) MCMAHON A P.  
 XX (TABI/) TABIN C J.  
 XX Ingham PW, McMahon AP, Tabin CJ;  
 XX WPI; 2003-803151/75.  
 XX N-PSDB; AAD62097.  
 XX  
 XX Modulating cell growth, differentiation or survival, for treating  
 XX neurodegenerative diseases, such as Alzheimer's or Parkinson's disease,  
 XX comprises contacting the cell with a hedgehog polypeptide.  
 XX  
 XX Claim 5; Page 82-83; Opp: English.  
 XX  
 CC The present invention relates to a novel method for modulating growth,  
 CC differentiation or survival of a cell. The method involves contacting the  
 CC cell with a hedgehog polypeptide such as Desert hedgehog (Dhh), Sonic  
 CC hedgehog (shh) and Indian hedgehog (Ihh). The method is used to induce a  
 CC cell to differentiate to a neuronal cell phenotype. It is used to  
 CC modulate skeletogenesis. The method is used to treat a degenerative  
 CC disorders of the nervous system such as neuromuscular, autonomic or  
 CC central nervous system disorders (e.g., Alzheimer's disease, Parkinson's  
 CC disease, amyotrophic lateral sclerosis, Huntington's disease, multiple  
 CC sclerosis, Pick's disease, neuronal degeneration associated with a  
 CC natural aging process and neuronal damage resulting from trauma and  
 CC neuronal damage resulting from anoxia-ischaemia. The invention is also  
 CC used for antisense gene therapy. The present sequence is human Shh  
 CC protein  
 XX  
 SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 7; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-217;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGGFGKRRHPKLTPLAYKQIPNVAEKLTCASG 60  
 DB 1 MLLARCLLLVSSLLVCSGLACGPGGFGKRRHPKLTPLAYKQIPNVAEKLTCASG 60  
 QY 61 RYEGKISNSRERFELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQWPGV 120  
 DB 61 RYEGKISNSRERFELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAI SVMNQWPGV 120  
 QY 121 KLRTVEGWDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
 DB 121 KLRTVEGWDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
 QY 181 IHCSVKAENSVAAKSGGCGFPASATVHLPGQGTCLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
 DB 181 IHCSVKAENSVAAKSGGCGFPASATVHLPGQGTCLVKOLSPGDRVLAADDQGRLLYSDFLT 240

Db 181 IHCSVKAENSVAAKSGCGPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 QY 241 FLDRDDGAKKVFVYIETREPRERLILLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFVYIETREPRERLILLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
 QY 301 PRALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
 QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGNMAVKSXSRGAGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGNMAVKSXSRGAGGAREGA 475

## RESULT 23

ADD71383  
 ID ADD71383 standard; protein; 475 AA.

XX AC ADD71383;

XX AC 29-JAN-2004 (first entry)

XX DT Human sonic hedgehog (shh).

XX DE hedgehog polypeptide; tissue array generation; tissue array maintenance;  
 KW human; sonic hedgehog; shh.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT Misc-difference 463 /note= "Unknown"

XX FT US2003190696-A1.

XX PN 09-OCT-2003.

XX PD 13-DEC-2000; 2000US-00736476.

XX PF 30-DEC-1993; 93US-00176427.

XX PR 14-DEC-1994; 94US-00356060.

XX PR 04-MAY-1995; 95US-00435093.

XX PR 05-JUN-1995; 95US-00460900.

XX PA (HARD ) HARVARD COLLEGE.

XX PI Ingham PW, McMahon AP, Tabin CJ, Buncrot DA, Marti-Gorostiza E;

XX PN WPI; 2003-831623/77.

XX DR N-PSDB; ADD71376.

XX FT New nucleic acid encoding a hedgehog polypeptide having an amino acid  
 FT sequence identical or homologous to a vertebrate hedgehog protein, useful  
 FT for generating or maintaining an array of different vertebrate tissue in  
 FT vitro and in vivo.

XX PS Claim 12; SEQ ID NO 13; 118pp; English.

XX CC The invention describes an isolated nucleic acid encoding a hedgehog  
 CC polypeptide having an amino acid sequence identical or homologous to a  
 CC vertebrate hedgehog protein or its portion and not identical to a fully  
 CC defined 471-bp sequence. The nucleic acid is useful for generating and/or  
 CC maintaining an array of different vertebrate tissue both in vitro and in  
 CC vivo. This is the amino acid sequence of human sonic hedgehog (shh).

XX SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 7; Length 475;

Best Local Similarity 100.0%; Pred. No. 2,1e-217; Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLILARCLLLVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
 Db 1 MLILARCLLLVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
 QY 61 RYEGKISRNSERPKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALAISVNNQWPGV 120  
 Db 61 RYEGKISRNSERPKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALAISVNNQWPGV 120  
 QY 121 KLRVTEGWDEDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
 Db 121 KLRVTEGWDEDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
 QY 181 IHCSVKAENSVAAKSGCGPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 181 IHCSVKAENSVAAKSGCGPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 QY 241 FLDRDDGAKKVFVYIETREPRERLILLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
 Db 241 FLDRDDGAKKVFVYIETREPRERLILLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
 QY 301 PRALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
 Db 301 PRALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
 QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 Db 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGNMAVKSXSRGAGGAREGA 475  
 Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGNMAVKSXSRGAGGAREGA 475

## RESULT 24

AAB85085  
 ID AAB85085 standard; protein; 462 AA.

XX AC AAB85085;

XX DT 22-AUG-2001 (first entry)

XX DE Human sonic hedgehog (Shh) polypeptide.

XX DE Insulin; hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh;  
 KW desert hedgehog; Dhh; diabetes; pancreatic beta-cell; PBC; IDX-1;  
 KW neogenesis; hyperinsulinemia.

XX OS Homo sapiens.

XX PN WO200141786-A1.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000WO-US033575.

XX PR 10-DEC-1999; 99US-0170282P.

XX PA (GEO ) GEN HOSPITAL CORP.

XX PI Habener JF, Thomas MK;

XX DR WPI; 2001-381492/40.

XX DR N-PSDB; AAF84005.

XX PT Treating deficiency of insulin, IDX-1 or pancreatic beta cells in a  
 PT patient by, administering a hedgehog protein, nucleic acid encoding the  
 PT protein or cells expressing the protein.

XX PS Disclosure; Fig 1B; 63pp; English.

CC The invention relates to a method of treating deficiency of insulin, that  
CC involves administering a hedgehog protein or nucleic acid encoding the  
CC hedgehog protein. The hedgehog proteins that can be used in the method  
CC are selected from sonic hedgehog (Shh), indian hedgehog (Ihh) and desert  
CC hedgehog (Dhh). The method is useful for treating deficiency of insulin  
CC in a patient afflicted with diabetes, by stimulating insulin production  
CC in pancreatic beta-cells (PBC). It is also used to treat deficiency of IDH  
CC -1 in a patient by stimulating IDH-1 production in PBC. The hedgehog  
CC protein is useful for modulating IDH-1 gene expression or its protein in  
CC PBC. This is used to treat deficiency of PBC in a patient, by stimulating  
CC neogenesis form beta-cell pancreatic ductal precursor cells. Inhibitors  
CC of the hedgehog proteins are useful for suppressing secretion of insulin  
CC in a patient afflicted with hyperinsulinemia. The present sequence  
CC represents a human Shh polypeptide  
XX  
XX Sequence 462 AA;

Query Match 97.5%; Score 2407; DB 4; Length 462;  
Best Local Similarity 100.0%; Pred. No. 6.6e-212;  
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MLLARCLLLVSLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVSLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISSVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISSVMNQPGV 120  
Qy 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGPDWVYYESKAH 180  
Db 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGPDWVYYESKAH 180  
Qy 181 IHCSVKAENSVAAGSGGCGPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCGPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSTATGPEASSGGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSTATGPEASSGGPPSGGALG 300  
Qy 301 PRALFASRVPGQRYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVPGQRYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
Db 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
Qy 421 ADAPGAGATAGTHWYSQLLYQIGTWLLDSEALHPLGMVAKSS 462  
Db 421 ADAPGAGATAGTHWYSQLLYQIGTWLLDSEALHPLGMVAKSS 462

RESULT 25  
ADA09286  
ID ADA09286 standard; protein; 462 AA.  
XX  
XX AC ADA09286;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Human Sonic Hedgehog protein (Shh).  
XX  
XX Human; hedgehog protein; desert; indian; sonic; Dhh; Ihh; Shh;  
XX cystostatic; gene therapy; carcinogenesis; gastric; colonic; cancer;  
XX Gastrointestinal tract; colonic adenomatous polyp;  
XX invasive adenocarcinoma; intestinal adenoma; desmoid tumour.  
XX  
XX Homo sapiens.  
XX  
XX WO2003070265-A2.

PD 28-AUG-2003.  
XX  
XX 20-FEB-2003; 2003WO-NL000127.  
XX  
XX 20-FEB-2003; 2002EP-00075690.  
XX  
XX (UYAM-) UNIV AMSTERDAM ACAD ZIEKENHUIS BIJ VAN.  
XX  
XX Van Den Brink GR, Peppelenbosch MP, Hardwick JCH, Van Deventer SJH;  
XX WPI; 2003-697568/66.  
XX  
XX Use of a Hedgehog protein for treating Hedgehog protein deficiency in the  
XX gastrointestinal tract, or for preventing gastric or colonic cancer,  
XX colonic adenomatous polyps, invasive adenocarcinomas, small intestinal  
XX adenomas.  
XX  
XX Disclosure; Page 47-49; 67pp; English.  
XX  
XX The invention relates to treating a deficiency of a Hedgehog protein in  
XX the gastrointestinal (GI) tract. The method of the invention comprises  
XX providing a source of Hedgehog protein to the GI tract of a subject  
XX suffering from the deficiency of this protein. Hedgehog proteins include  
XX Desert (Dhh), Indian (Ihh) and Sonic (Shh). The treatment of the  
XX deficiency of a Hedgehog protein in the GI tract is useful for the  
XX prophylaxis of carcinogenesis in the GI tract, gastric or colonic cancer,  
XX and for the treatment of a GI tract carcinoma, particularly gastric or  
XX colonic cancer. The source Hedgehog protein may also be administered to  
XX prevent or reverse colonic adenomatous polyps, invasive adenocarcinomas,  
XX small intestinal adenomas and cancers, and desmoid tumors. Loss of Indian  
XX hedgehog (Ihh) expression was evident in the earliest recognisable stage  
XX in the sequence of events that is thought to progress to colon cancer.  
XX Ihh was found to be not expressed in malignant colonic epithelial cells.  
XX The colon cancer cell HT-29 will only express Ihh protein after  
XX differentiation with butyrate. The in vitro differentiation of HT-29  
XX cells with recombinant hedgehog (Ihh) protein was determined. Forty-eight  
XX hour treatment of HT-29 cells induces Willin expression to a similar  
XX extent as after treatment with 5 mM butyrate. Data show that exogenous Ihh  
XX protein was sufficient to restore differentiation of colon carcinoma  
XX cells. Hedgehog proteins may also be used in gene therapy techniques. The  
XX current sequence represents the human Sonic hedgehog protein (Shh).  
XX  
XX Sequence 462 AA;

Query Match 97.5%; Score 2407; DB 6; Length 462;  
Best Local Similarity 100.0%; Pred. No. 6.6e-212;  
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MLLARCLLLVSLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVSLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISSVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISSVMNQPGV 120  
Qy 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGPDWVYYESKAH 180  
Db 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGPDWVYYESKAH 180  
Qy 181 IHCSVKAENSVAAGSGGCGPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCGPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSTATGPEASSGGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSTATGPEASSGGPPSGGALG 300  
Qy 301 PRALFASRVPGQRYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVPGQRYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420

Db 361 ASYAVIEEHSWAHRAFAFRLAHALLAALAPARTDRGGSGGGGRVALTAPGA 420

Qy 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGMAVKS 462

Db 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGMAVKS 462

RESULT 26

AA85738

ID AAB85738 standard; protein; 463 AA.

XX AC AAB85738;

XX DT 29-OCT-2001 (first entry)

XX DE Human sonic hedgehog (Shh) polypeptide.

XX KW Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh; Dhh;

XX KW desert hedgehog; cell differentiation; human.

XX OS Homo sapiens.

XX PN US6271363-B1.

XX PD 07-AUG-2001.

XX PF 20-OCT-1997; 97US-00954698.

XX PR 30-DEC-1993; 93US-00176427.

XX PR 14-DEC-1994; 94US-00356060.

XX PR 04-MAY-1995; 95US-00435092.

XX PR 05-JUN-1995; 95US-00462386.

XX PA (HARD ) HARVARD COLLEGE.

XX PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.

XX PI Ingham PW, McMahon AP, Tabin CJ;

XX XN WPI; 2001-456723/49.

XX DR N-PSDB; AAH76112.

XX PT Novel nucleic acid encoding a hedgehog polypeptide, used to produce the

XX PT polypeptide, which is used to promote proliferation, survival, and/or

XX PT differentiation of neuronal and mesodermal tissue.

XX PS Claim 1; Col 133-136; 118pp; English.

XX CC The invention relates to nucleic acids encoding hedgehog proteins

XX CC selected from sonic hedgehog (Shh), indian hedgehog (Ihh), desert

XX CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the

XX CC formation of ordered spatial arrangements of differentiated tissue in

XX CC vertebrates. The nucleic acid sequences are useful for producing hedgehog

XX CC proteins, used for promoting differentiation of, or survival of

XX CC differentiated, neuronal cells, and for promoting proliferation, survival

XX CC or differentiation of mesenchymal, endodermal or ectodermal tissue,

XX CC particularly chondrocytes, or testicular germ line cells. The present

XX CC sequence represents a human Shh polypeptide

XX SQ Sequence 463 AA;

Query Match 97.5%; Score 2407; DB 4; Length 463;

Best Local Similarity 100.0%; Pred. No. 6.7e-212;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLVVSSLLVCSGLACGPGRGKRRHPKKLTPLAYKQIPNVAEKTGASG 60

Db 1 MLLARCLLVVSSLLVCSGLACGPGRGKRRHPKKLTPLAYKQIPNVAEKTGASG 60

Qy 61 RYEGKISRNSERFKELTNPNDIIFKDEENTGADRLMTQCKLNALAISSVNNQPGV 120

Db 61 RYEGKISRNSERFKELTNPNDIIFKDEENTGADRLMTQCKLNALAISSVNNQPGV 120

Qy 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGLMARLAVEAGFDWVYYSKAH 180

Db 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGLMARLAVEAGFDWVYYSKAH 180

Qy 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDGRLLYSDFLT 240

Db 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGTKLVKDLSPGDRVLAADDGRLLYSDFLT 240

Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300

Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300

Qy 301 PRALFASRVRRGQVYVVAERDGRLLPAVHVSVTLSSEAAAGAYAPLTAGTILLINRVL 360

Db 301 PRALFASRVRRGQVYVVAERDGRLLPAVHVSVTLSSEAAAGAYAPLTAGTILLINRVL 360

Qy 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDRGGSGGGGRVALTAPGA 420

Db 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDRGGSGGGGRVALTAPGA 420

Qy 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGMAVKS 462

Db 421 ADAPGAGATAGIHWSYQLLYQIGTWLDDSEALHPLGMAVKS 462

RESULT 27

AAW48736

ID AAW48736 standard; protein; 462 AA.

XX AC AAW48736;

XX DT 14-SEP-1998 (first entry)

XX DE Human mutated sonic hedgehog (SHH) protein (Met-114Ile).

XX KW Sonic hedgehog; SHH protein; tumorigenesis; oncogenesis;

XX KW basal cell carcinoma; breast cancer; medulloblastoma; tumour;

XX KW cell proliferation; cell differentiation; wound healing; diagnosis;

XX KW therapy; human.

XX OS Homo sapiens.

XX XN WO9821227-A1.

XX PD 22-MAY-1998.

XX PF 12-NOV-1997; 97WO-US020227.

XX PR 13-NOV-1996; 96US-00748591.

XX PA (REGC ) UNIV CALIFORNIA.

XX PI Epstein E, Hu Z, Bonifas J;

XX DR WPI; 1998-297857/26.

XX DR N-PSDB; AAV18404.

XX PT New nucleic acid encoding oncogenic human hedgehog protein - useful for,

XX PT e.g. treatment and diagnosis of cancer and diseases involving cell

XX PT proliferation or differentiation.

XX PS Claim 6; Page 32-33; 47pp; English.

XX CC This human sonic hedgehog (SHH) protein carries a spontaneously occurring

XX CC mutation (M141 substitution) that is associated with oncogenic

XX CC transformation of human cells. The amino acid substitution results from a

XX CC mutation of the SHH gene (see AAV18404). This mutation was identified in

XX CC a basal cell carcinoma. A second basal cell carcinoma, as well as a

XX CC medullocarcinoma and a breast cancer were shown to carry an independently

XX CC arising different mutation (see AAW48735). Purified oncogenic HH

XX CC proteins, and methods for producing the proteins using mammalian

XX CC (preferably human) host cells are claimed. The HH proteins can be used;

XX CC in functional mapping; therapeutically for modulating cell proliferation

CC and differentiation, e.g. in cases of wound healing or any of a very wide  
CC range of conditions involving reduced hedgehog signalling such as bone  
CC formation, hyperproliferative or hypoproliferative skin disease, growth  
CC of hair, neurodegeneration (Alzheimer's disease etc.) or autonomic  
CC disorders of the peripheral nervous system, e.g. cardiac arrhythmia; to  
CC study physiological pathways; and to screen for drugs that modulate HH  
CC activity and expression. HH proteins are also useful in vitro for  
CC sustaining reproduction of neural progenitor cells. Detecting presence of  
CC oncogenic HH mutations is used to characterise the phenotype of a tumour;  
CC similar analyses can be done at the protein level using antibodies, which  
CC may also be used to treat human cancers  
XX  
SQ Sequence 462 AA;

Query Match 97.3%; Score 2403; DB 2; Length 462;  
Best Local Similarity 99.8%; Pred. No. 1.5e-211;  
Matches 461; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLLTPLAYKQFIPNVAEKTIGASG 60  
Qy 61 RYEGKISRNSERFKELTPNYNPDIIIFKDEENTGADRLMTQCKDKLNALAISSVMNQWPGV 120  
Db 61 RYEGKISRNSERFKELTPNYNPDIIIFKDEENTGADRLMTQCKDKLNALAISSVMNQWPGV 120  
Qy 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 180  
Db 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 180  
Qy 181 IHCSVKAENSVAAGSGCFPGSATVHLEQGKTLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGCFPGSATVHLEQGKTLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
Qy 241 FLDRDDGAKVYVIETREPRERULLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKVYVIETREPRERULLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Qy 301 PRALFASVRPQGRVYVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPQGRVYVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 462

RESULT 28  
AAW48735  
ID AAW48735 standard; protein; 462 AA.  
XX  
AC AAW48735;  
XX  
DT 14-SEP-1998 (first entry)  
XX  
DE Human mutated sonic hedgehog (SHH) protein (His133-Tyr).  
XX  
KW Sonic hedgehog; SHH protein; tumorigenesis; oncogenesis;  
KW basal cell carcinoma; breast cancer; medulloblastoma; tumour;  
KW cell proliferation; cell differentiation; wound healing; diagnosis;  
KW therapy; human.  
XX  
OS Homo sapiens.  
XX  
FN WO9821227-A1.  
XX  
PD 22-MAY-1998.  
XX  
XX 12-NOV-1997; 97WO-05020227.

XX 13-NOV-1996; 96US-00748591.  
XX (REGC ) UNIV CALIFORNIA.  
XX  
XX Epstein E, Hu Z, Bonifas J;  
XX WPI; 1998-297857/26.  
XX N-PSDB; AAV18403.  
XX New nucleic acid encoding oncogenic human hedgehog protein - useful for,  
XX e.g. treatment and diagnosis of cancer and diseases involving cell  
XX proliferation or differentiation.  
XX  
XX Claim 6; Page 28-29; 47pp; English.  
XX  
XX This human sonic hedgehog (SHH) protein carries a spontaneously occurring  
XX mutation (His133 substitution) that is associated with oncogenic  
XX transformation of human cells. The amino acid substitution results from a  
XX mutation of the SHH gene (see AAV18403). This mutation was identified in  
XX 3 independently arising tumours (a basal cell carcinoma  
XX medullocarcinoma, and breast cancer). A second basal cell carcinoma was  
XX shown to carry a different mutation (see AAW48736). Purified oncogenic HH  
XX proteins, and methods for producing the proteins using mammalian  
XX (preferably human) host cells are claimed. The HH proteins can be used:  
XX in functional mapping; therapeutically for modulating cell proliferation  
XX and differentiation, e.g. in cases of wound healing or any of a very wide  
XX range of conditions involving reduced hedgehog signalling such as bone  
XX formation, hyperproliferative or hypoproliferative skin disease, growth  
XX of hair, neurodegeneration (Alzheimer's disease etc.) or autonomic  
XX disorders of the peripheral nervous system, e.g. cardiac arrhythmia; to  
XX study physiological pathways; and to screen for drugs that modulate HH  
XX activity and expression. HH proteins are also useful in vitro for  
XX sustaining reproduction of neural progenitor cells. Detecting presence of  
XX oncogenic HH mutations is used to characterise the phenotype of a tumour;  
XX similar analyses can be done at the protein level using antibodies, which  
XX may also be used to treat human cancers  
XX

SQ Sequence 462 AA;

Query Match 97.2%; Score 2401; DB 2; Length 462;  
Best Local Similarity 99.8%; Pred. No. 2.4e-211;  
Matches 461; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLLTPLAYKQFIPNVAEKTIGASG 60  
Qy 61 RYEGKISRNSERFKELTPNYNPDIIIFKDEENTGADRLMTQCKDKLNALAISSVMNQWPGV 120  
Db 61 RYEGKISRNSERFKELTPNYNPDIIIFKDEENTGADRLMTQCKDKLNALAISSVMNQWPGV 120  
Qy 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 180  
Db 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 180  
Qy 181 IHCSVKAENSVAAGSGCFPGSATVHLEQGKTLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGCFPGSATVHLEQGKTLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
Qy 241 FLDRDDGAKVYVIETREPRERULLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKVYVIETREPRERULLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Qy 301 PRALFASVRPQGRVYVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPQGRVYVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Qy 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
Db 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 462

Db	421	ADAPGAGATAGIHWYSQLYQIGTWLLDSEALHPLGMVAKSS	462
Db	AAW61488	standard; protein; 437 AA.	
XX	AAW61488;		
XX	AC		
XX	25-MAR-2003 (revised)		
DT	20-OCT-1998 (first entry)		
DT	XX		
XX	Mouse sonic hedgehog (shh) protein.		
DE	XX		
XX	Hedgehog polypeptide; neuronal cell proliferation; zebrafish; shh;		
KW	cholesterol biosynthesis; pituitary gland gene expression; mouse;		
KW	tiggy-winkle hedgehog protein; twih; sonic hedgehog; major limb trauma;		
KW	neuronal degeneration; nerve-sparing agent; Smith-Lemli-Optiz syndrome.		
XX	XX		
OS	Mus sp.		
XX	XX		
PN	WO9830576-A1.		
XX	XX		
PD	16-JUL-1998.		
XX	07-OCT-1997; 97WO-US015753.		
PF	XX		
XX	07-OCT-1996; 96US-00729743.		
PR	02-OCT-1997; 97US-0061323P.		
PR	XX		
XX	(UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.		
PA	XX		
XX	Beachy'PA, Porter JA;		
PI	XX		
XX	WPI; 1998-399053/34.		
DR	XX		
XX	New hedgehog-derived poly:peptide(s) - used to develop products for		
PT	modulating proliferation or differentiation of neuronal cells,		
PT	cholesterol biosynthesis or transport or expression of pituitary gland		
PT	gene(s).		
XX	XX		
XX	Claim 13; Page 153-154; 210pp; English.		
PS	XX		
XX	This represents a mouse sonic hedgehog (shh) protein sequence. The		
CC	invention provides methods and compounds for modulating proliferation or		
CC	differentiation of neuronal cells, cholesterol biosynthesis or transport		
CC	or expression of pituitary gland genes. The method for affecting		
CC	cholesterol biosynthesis or transport in a cell comprises contacting a		
CC	cell with a compound that affects hedgehog, thereby affecting cholesterol		
CC	biosynthesis or transport. The methods for inhibiting the neural inducing		
CC	activity of a hedgehog polypeptide in cells, and for inducing pituitary		
CC	gland gene expression utilises sequences selected from a zebrafish tiggy-		
CC	winkle hedgehog (twih) protein (AAW61485), a zebrafish sonic hedgehog		
CC	(shh) protein (AAW61486), a chicken shh protein (AAW61487) or a mouse shh		
CC	protein (AAW61488). The products and methods provide for compounds which		
CC	can affect hedgehog activity. They can be used for treating disorders		
CC	which arise from neuronal degeneration or abnormal function. They can		
CC	also be used as nerve-sparing agents or in restoring or promoting		
CC	appropriate patterning during the healing of major limb trauma. They can		
CC	also be used for treating Smith-Lemli-Optiz syndrome. The products can		
CC	also be used for detection and diagnosis. (Updated on 25-MAR-2003 to		
CC	correct PA field.)		
XX	XX		
SQ	Sequence 437 AA;		
XX	Query Match	83.8%; Score 2068; DB 2; Length 437;	
XX	Best Local Similarity	87.4%; Pred. No. 8.3e-181;	
XX	Matches 404; Conservative	8; Mismatches 24; Indels 26; Gaps 3;	
QY	1	MLLARCILLVLSGLVCSLACGPGRGKRRHPKMLPLAYKQFIPNVAEKTIGASG	60
Db	2	LLLARCFLVILASSLVLCGLACGPGRGKRRHPKMLPLAYKQFIPNVAEKTIGASG	61

QY	61	RYEGKISRNSERPKELTPNNPDIIFKDEENTGADLMTQRCCKLNALAI SVMQWPGV	120
Db	62	RYEGKISRNSERPKELTPNNPDIIFKDEENTGADLMTQRCCKLNALAI SVMQWPGV	121
QY	121	KLRTVEGWEDDGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYESKAH	180
Db	122	KLRTVEGWEDDGHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYESKAH	181
QY	181	IHCSTVAENSVAAGSGCPCPGSATVHLEOGGTYKLVKDLSPGDRVLAADDQGRLLYSDFLT	240
Db	182	IHCSTVAENSVAAGSGCPCPGSATVHLEOGGTYKLVKDLSPGDRVLAADDQGRLLYSDFLT	241
QY	241	FLDRDDGAKKVFVVIETREPRERLLITAAHLLFVAPHNDSATCEPEASSGSGPPSGGALG	300
Db	242	FLDRDEGAKKVFVVIETLEPRERLLITAAHLLFVAPHND-----SGPTPG	286
QY	301	PRALFASRVPRGQVVVVAERDGRLLPAVHSVTLSSEAGAYAPLTAQGTILINRVL	360
Db	287	PSALFASRVPRGQVVVVAERDGRLLPAVHSVTLSSEAGAYAPLTAHGTILINRVL	346
QY	361	ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGGDSGGGGRGGRVALTAPGA	420
Db	347	ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGSGIP-AAQSA	395
QY	421	ADAPGAGATAGIHWYSQLYQIGTWLLDSEALHPLGMVAKSS	462
Db	396	TEARGAEPITAGIHWYSQLYHIGTWLLDSEALHPLGMVAKSS	437
XX	RESULT 30		
XX	AAB85086		
ID	AAB85086	standard; protein; 437 AA.	
AC	AAB85086;		
XX	XX		
DT	22-AUG-2001 (first entry)		
XX	XX		
DE	Mouse sonic hedgehog (Shh) polypeptide.		
XX	XX		
KW	Insulin; hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh;		
KW	desert hedgehog; Dhh; diabetes; pancreatic beta-cell; PBC; IDX-1;		
KW	neogenesis; hyperinsulinemia.		
XX	XX		
OS	Mus musculus.		
XX	XX		
PN	WO200141786-A1.		
XX	XX		
PD	14-JUN-2001.		
XX	XX		
PF	08-DEC-2000; 2000WO-US033575.		
XX	XX		
PR	10-DEC-1999; 99US-0170282P.		
XX	XX		
PA	(GEHO ) GEN HOSPITAL CORP.		
PI	Habener JF, Thomas MK;		
XX	XX		
DR	WPI; 2001-381492/40.		
XX	N-FSDB; AAF84006.		
XX	XX		
PT	Treating deficiency of insulin, IDX-1 or pancreatic beta cells in a		
PT	patient by, administering a hedgehog protein, nucleic acid encoding the		
PT	protein or cells expressing the protein.		
XX	XX		
PS	Disclosure; Fig 2B; 63pp; English.		
XX	XX		
CC	The invention relates to a method of treating deficiency of insulin, that		
CC	involves administering a hedgehog protein or nucleic acid encoding the		
CC	hedgehog protein. The hedgehog proteins that can be used in the method		
CC	are selected from sonic hedgehog (Shh), indian hedgehog (Ihh) and desert		
CC	hedgehog (Dhh). The method is useful for treating deficiency of insulin		
CC	in a patient afflicted with diabetes, by stimulating insulin production		

CC in pancreatic beta-cells (PBC). It is also used to treat deficiency of IDX  
CC -1 in a patient, by stimulating IDX-1 production in PBC. The hedgehog  
CC protein is useful for modulating IDX-1 gene expression or its protein in  
CC PBC. This is used to treat deficiency of PBC in a patient, by stimulating  
CC neogenesis form beta-cell pancreatic ductal precursor cells. Inhibitors  
CC of the hedgehog proteins are useful for suppressing secretion of insulin  
CC in a patient afflicted with hyperinsulinemia. The present sequence  
CC represents a mouse Shh polypeptide  
XX  
SQ Sequence 437 AA;

Query Match 83.8%; Score 2068; DB 4; Length 437;  
Best Local Similarity 87.4%; Pred. No. 8.3e-181;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;  
Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTIGASG 60  
Db 2 LLLARCFVLIVASSLLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTIGASG 61  
Qy 61 RYEGKISRNSERFKELTPNPNPDIIFKDEENTGADRLMTORCKDKNALAISVMNQWPGV 120  
Db 62 RYEGKITRNSERFKELTPNPNPDIIFKDEENTGADRLMTORCKDKNALAISVMNQWPGV 121  
Qy 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGDFWVYYESKAH 180  
Db 122 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGDFWVYYESKAH 181  
Qy 181 IHCSVKAENSVAKSGCGFPGSATVHLEOGGKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAKSGCGFPGSATVHLEOGGKLVKDLSPGDRVLAADDGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVFYVITREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVITREPRERLLTAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASRVPCQVYVVAERDGRLLPAAVHVSVTLSERAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPCQVYVVAERGDRLLPAAVHVSVTLSERAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTDGGSDGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTD-----GGGGGSIP-AAQSA 395  
Qy 421 ADAGAGATAGIHWYSQLLYOIGTWLIDSEALHPLGMVAKSS 462  
Db 396 TEARGAETAGIHWYSQLLYHIGTWLIDSETMPLGMVAKSS 437

RESULT 31  
AAE09292  
ID AAE09292 standard; protein; 437 AA.  
XX  
AC AAE09292;  
XX  
DT 19-NOV-2001 (first entry)  
XX  
DE Mouse sonic hedgehog (Shh) protein.  
XX  
KW Fibroblast growth factor-4; FGF-4; neuronal formation; dopamine; DA;  
KW serotonin; 5-hydroxytryptamine; 5HT; neuroprogenitor cell; depression;  
KW Sonic hedgehog; Shh; food intake disorder; stress response; vulnerability;  
KW sexual activity; cardiovascular function; temperature regulation; pain;  
KW obsessive-compulsive behaviour; violent aggressive behaviour; anorexia;  
KW reward-associated behaviour; Parkinson's disease; sensory neglect;  
KW drug addiction; lesion; trauma; illness; tremor; rigidity; neuroleptic;  
KW postural abnormality; akinesia; aphasia; schizophrenia; adipsia; mouse;  
XX  
XX Mus musculus.  
OS  
XX  
PN US6277820-B1.  
XX  
PD 21-AUG-2001.

XX 09-APR-1998; 98US-00057860.  
XX 09-APR-1998; 98US-00057860.  
XX (GETH ) GENE TECH INC.  
XX Rosenthal A, Hynes MA, Ye W;  
XX WPI; 2001-540411/60.  
XX N-PSDB; AAD16025.  
XX Forming dopaminergic neurons for treating disorders due to abnormalities  
XX in postural reflex regulation, involves contacting neuroprogenitor cells  
XX with fibroblast growth factor-8 and sonic hedgehog polypeptide.  
XX  
XX Claim 1; Fig 11; 40pp; English.  
XX  
XX The present invention relates to neuronal formation and methods of  
XX treating diseases characterised by abnormalities in the activity of  
XX dopaminergic (DA) and serotonergic (5HT) neurons. The invention also  
XX relates to a method of forming serotonergic neurons in vitro by  
XX contacting neuroprogenitor cells to an effective amount of native  
XX sequence, variants and functional fragments of fibroblast growth factor-8  
XX (FGF-8), FGF-4 and Sonic hedgehog (Shh). The method is used to treat  
XX disorders relating to food intake, hormone secretion, stress response,  
XX pain and immune function, sexual activity, cardiovascular function and  
XX temperature regulation, in particular depression, proclivity to suicide,  
XX obsessive-compulsive behaviour, violent aggressive behaviour and  
XX anorexia/bulimia. The method and the composition are useful for forming  
XX dopaminergic neurons by stimulating differentiation of neuroprogenitor  
XX cells into dopaminergic neurons which is useful for treating disorders  
XX characterised by abnormalities in the regulation of postural reflexes,  
XX movement and reward-associated behaviours including Parkinson's disease,  
XX schizophrenia, drug addiction, lesions due to trauma or other illness  
XX resulting in Parkinson-like conditions such as resting tremor, rigidity,  
XX akinesia and postural abnormality, including akinesia, adipsia, aphasia  
XX and sensory neglect. The present sequence is mouse sonic hedgehog (Shh)  
XX protein  
XX  
SQ Sequence 437 AA;  
Query Match 83.8%; Score 2068; DB 4; Length 437;  
Best Local Similarity 87.4%; Pred. No. 8.3e-181;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;  
Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTIGASG 60  
Db 2 LLLARCFVLIVASSLLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTIGASG 61  
Qy 61 RYEGKISRNSERFKELTPNPNPDIIFKDEENTGADRLMTORCKDKNALAISVMNQWPGV 120  
Db 62 RYEGKITRNSERFKELTPNPNPDIIFKDEENTGADRLMTORCKDKNALAISVMNQWPGV 121  
Qy 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGDFWVYYESKAH 180  
Db 122 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGDFWVYYESKAH 181  
Qy 181 IHCSVKAENSVAKSGCGFPGSATVHLEOGGKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAKSGCGFPGSATVHLEOGGKLVKDLSPGDRVLAADDGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVFYVITREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVITREPRERLLTAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASRVPCQVYVVAERDGRLLPAAVHVSVTLSERAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPCQVYVVAERGDRLLPAAVHVSVTLSERAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTDGGSDGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPFRLAHLAALAPARTD-----GGGGGSIP-AAQSA 395

QY 421 ADAPGAGATAGIHWYSQLYQIGTWLSDSEALHPLGMVAVKSS 462  
 Db 396 TEARGAEPTAGIHWYSQLYHIGTWLSDSETMPLGMVAVKSS 437

RESULT 32  
 ID AAY97559  
 XX AAY97559 standard; protein; 437 AA.  
 AC AAY97559;  
 XX  
 DT 05-APR-2001 (first entry)  
 XX  
 DE Mouse sonic hedgehog protein sequence.  
 XX  
 KW Mouse; Hedgehog signalling pathway; inhibitor; autoimmune disorder;  
 KW epithelial cell hyperplasia; fibrosis; inflammation; adenocarcinoma;  
 KW immune disorder; cancer; thyroiditis; insulinitis; multiple sclerosis;  
 KW iridocyclitis; uveitis; orchitis; hepatitis; Addison's disease; asthma;  
 KW myasthenia gravis; rheumatoid arthritis; lupus erythematosus; emphysema;  
 KW adult respiratory distress syndrome; chronic bronchitis; atelectasis;  
 KW silicosis; hypersensitivity pneumonitis; idiopathic pulmonary fibrosis;  
 KW pneumonia; pleural fibrosis; atherosclerosis; myocardial infarction;  
 KW Gastrointestinal tract disorder; hepatic disease; Alzheimer's disease;  
 KW Wnt signalling pathway; BMP signalling pathway; sonic hedgehog; SHH.  
 XX  
 OS Mus musculus.  
 XX  
 XX WO200074706-A1.  
 XX  
 PN 14-DEC-2000.  
 XX  
 PD 05-JUN-2000; 2000WO-GB002191.  
 XX  
 PF 08-JUN-1999; 99GB-00013350.  
 PR 16-SEP-1999; 99GB-00021953.  
 XX  
 XX (LORA-) LORANTIS LTD.  
 XX  
 XX Lamb JR, Hoynes GF, Dallman MJ;  
 PI WPI; 2001-061652/07.  
 DR N-PSDB; AAA37897.  
 XX  
 XX Use of an inhibitor of a Hedgehog signalling pathway in preparation of a  
 PT medicament for treating epithelial cell hyperplasia, inflammation, cancer  
 PT or an immune disorder.  
 XX  
 PS Disclosure; Page 70; 78pp; English.

This sequence is mouse sonic hedgehog (SHH), and is part of the Hedgehog signalling pathway. The invention relates to the use of an inhibitor of a Hedgehog signalling pathway, or an inhibitor of a target pathway of the Hedgehog signalling pathway, in the preparation of a medicament for treating epithelial cell hyperplasia, fibrosis of tissue, inflammation, cancer (especially adenocarcinoma) or an immune disorder. The immune disorder can be an autoimmune disorder such as thyroiditis, insulinitis, multiple sclerosis, iridocyclitis, uveitis, orchitis, hepatitis, Addison's disease, myasthenia gravis, rheumatoid arthritis or lupus erythematosus. Medicament containing the inhibitors is also useful for treating lung or kidney diseases such as adult respiratory distress syndrome, chronic obstructive airway disorders including asthma, emphysema and chronic bronchitis, atelectasis, occupational lung disease including silicosis, hypersensitivity diseases of lung such as hypersensitivity pneumonitis, idiopathic interstitial lung diseases such as idiopathic pulmonary fibrosis, pneumonia including interstitial pneumonia, desquamative interstitial pneumonia and acute interstitial pneumonia, and pleural fibrosis. They can also be used in the treatment of inflammation associated with atherosclerosis, e.g. myocardial infarction, diseases of the gastrointestinal tract, hepatic diseases and diseases of the central nervous system e.g. Alzheimer's disease. The hedgehog signalling pathway is that of Sonic hedgehog, Indian hedgehog or

CC Desert hedgehog, the target pathway of the hedgehog signalling pathway is  
 CC Wnt or BMP signalling pathway  
 XX  
 SQ Sequence 437 AA;

Query Match 83.8%; Score 2068; DB 4; Length 437;  
 Best Local Similarity 87.4%; Pred. No. 8.3e-181;  
 Matches 404; Conservative 6; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTLLGASG 60  
 Db 2 LLLARCLVILASSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTLLGASG 61  
 QY 61 RYEGKISRNSEKELTPNPNPDIIFKDBENTGADRLMTORCKDKLNALAISSVNNQWPGV 120  
 Db 62 RYEGKITRNSEKELTPNPNPDIIFKDBENTGADRLMTORCKDKLNALAISSVNNQWPGV 121  
 QY 121 KLRVTEGWDECHSESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 180  
 Db 122 KLRVTEGWDECHSESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 181  
 QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 241  
 QY 241 FLDRDDGAKKVPYVETREPERELLTAHLLFVAPHNDSATGPEASSSGSPSGGALG 300  
 Db 242 FLDRDEGAKKVPYVETLEPRERULLTAHLLFVAPHNDSATGPEASSSGSPSGGALG 300  
 QY 301 PRALFASRVREGQRYVYVAERDGRRLIPAAVHVSITLSEEAAGAYAPLTAGTTLINRVL 360  
 Db 287 PSALFASRVREGQRYVYVAERDGRRLIPAAVHVSITLSEEAAGAYAPLTAGTTLINRVL 346  
 QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGSDGGGGRGGRVALTAPGA 420  
 Db 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGSDGGGGRGGRVALTAPGA 395  
 QY 421 ADAPGAGATAGIHWYSQLYQIGTWLSDSEALHPLGMVAVKSS 462  
 Db 396 TEARGAEPTAGIHWYSQLYHIGTWLSDSETMPLGMVAVKSS 437

RESULT 33  
 AAO21264  
 ID AAO21264 standard; protein; 437 AA.  
 XX  
 AC AAO21264;  
 XX  
 DT 05-AUG-2002 (first entry)  
 XX  
 DE 437-mer wild-type sonic hedgehog protein.  
 XX  
 KW Sonic hedgehog protein; amino terminal protein; modulate differentiation;  
 KW signalling developmental process; neural plate; ligand; patched receptor;  
 KW wild-type.  
 XX  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 FH Cleavage-site 198..200  
 FT /label= GCF autoproteolytic site  
 FT /note= "This is an autoproteolytic site. Cleavage occurs  
 FT between the Gly-Cys residues"  
 XX  
 XX WO200224151-A2.  
 XX  
 XX 28-MAR-2002.  
 XX  
 XX 24-SEP-2001; 2001WO-US042438.  
 XX  
 XX 22-SEP-2000; 2000US-0235153P.  
 PR  
 XX (UYUO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 PA



QY 301 PRALFASRVPGQVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
 DB 287 PSALFASRVPGQVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAHGTILINRVL 346  
 QY 361 ASCYAVIEHSHWAHRAFPAPFLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
 DB 347 ASCYAVIEHSHWAHRAFPAPFLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
 QY 421 ADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462  
 DB 396 TEARGAETAGIHWSQLLYHIGTWLSDSETHPLGMVAKSS 437

RESULT 35  
 AAR77339  
 ID AAR77339 standard; protein; 437 AA.  
 XX AC AAR77339;  
 XX DT 08-MAR-1996 (first entry)  
 XX DE Mouse sonic hedgehog protein.  
 XX KW Mouse; sonic hedgehog protein; probe; primer; diagnostic;  
 KW nervous system disorder; gene therapy; antibody.  
 XX OS Mus musculus.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1..24 /note= "signal peptide"  
 FT Peptide 25..30 /note= "conserved sequence (AAR77345)"  
 FT Modified-site 279 /note= "N-linked glycosylation site"  
 FT WO9518856-A1.  
 XX PN 13-JUL-1995.  
 XX PP 30-DEC-1994; 94WO-US014992.  
 XX PR 30-DEC-1993; 93US-00176427.  
 XX PR 14-DEC-1994; 94US-00356060.  
 XX PA (HARD ) HARVARD COLLEGE.  
 XX PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
 XX PI Ingham FW, McMahon AP, Tabin CJ;  
 XX DR WPI; 1995-255060/33.  
 XX DR N-PSDB; AAQ91637.  
 XX PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful to  
 XX treat degenerative nervous system disorder(s) and in gene therapy.  
 XX PS Claim 17; Page 139-41; 210pp; English.  
 XX CC The sequence represents a mouse sonic hedgehog protein, homologous to a  
 CC Drosophila hedgehog protein (AAR77337), and is encoded by a cDNA isolated  
 CC from an 8.5-day post coitum mouse cDNA library. Probes and primers  
 CC derived from the sonic hedgehog gene may be used as diagnostic agents for  
 CC neuromuscular, autonomic or central nervous system disorders, and the  
 CC gene may also be used in gene therapy. Antibodies generated from the  
 CC protein may be used as therapeutic or research reagents  
 XX Sequence 437 AA;

Query Match 83.6%; Score 2065; DB 2; Length 437;  
 Best Local Similarity 87.2%; Pref. No. 1.6e-180;  
 Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLIARCLLVVSSLLVCSGLACGPGRGKRRHPKLTPLAYKOFIPNVAEKTIGASG 60  
 DB 2 LLLIARCLLVVSSLLVCSGLACGPGRGKRRHPKLTPLAYKOFIPNVAEKTIGASG 61  
 QY 61 RYEGKISRNERPKELTPNYPDIIPKDEENTGADRLMTORCKDKLNALAISSVMNWPVG 120  
 DB 62 RYEGKISRNERPKELTPNYPDIIPKDEENTGADRLMTORCKDKLNALAISSVMNWPVG 121  
 QY 121 KLRVTEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYESKAH 180  
 DB 122 RLRVTEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYESKAH 181  
 QY 181 IHC SVKAENSVAAKSGGCGFPSSATVHLEQGTQKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
 DB 182 IHC SVKAENSVAAKSGGCGFPSSATVHLEQGTQKLVKDLSPGDRVLAADDGRLLYSDFLT 241  
 QY 241 FLDRDDGAKVYVVIETREPRERILLTAHLFLVAPHNDSATGEPRASSGSGPPSGALG 300  
 DB 242 FLDRDEGAKVYVVIETREPRERILLTAHLFLVAPHND-----SGPTPG 286  
 QY 301 PRALFASRVPGQVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
 DB 287 PSALFASRVPGQVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAHGTILINRVL 346  
 QY 361 ASCYAVIEHSHWAHRAFPAPFLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
 DB 347 ASCYAVIEHSHWAHRAFPAPFLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
 QY 421 ADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462  
 DB 396 TEARGAETAGIHWSQLLYHIGTWLSDSETHPLGMVAKSS 437

RESULT 36  
 AAR94471  
 ID AAR94471 standard; protein; 437 AA.  
 XX AC AAR94471;  
 XX DT 29-APR-1999 (first entry)  
 XX DE Mouse Shh hedgehog protein sequence.  
 XX KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;  
 KW brain infarction; cerebral infarction; transient ischaemic attack;  
 KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;  
 KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.  
 XX OS Mus sp.  
 XX PN WO9900117-A2.  
 XX PD 07-JAN-1999.  
 XX PF 26-JUN-1998; 98WO-US013387.  
 XX PR 27-JUN-1997; 97US-00883656.  
 XX PA (ONTO-) ONTOGENY INC.  
 XX PI Mahanthappa NK;  
 XX DR WPI; 1999-095458/08.  
 XX DR N-PSDB; AAX16185.  
 XX PT Method for limiting damage to neurons caused by ischaemic or epoxic  
 PT conditions - is used for the treatment and prevention of e.g. cerebral  
 PT infarction, stroke and transient ischaemic attacks.  
 XX PS Disclosure; Page 68-70; 104pp; English.  
 XX CC A method has been developed for limiting the damage to neuronal cells by  
 CC ischaemic or epoxic conditions by administering a ptc (patched)

therapeutic agent to reduce cerebral infarct volume (CIV). Damage to neuronal cells can also be limited by administering a gene activation construct which recombines with the genomic hedgehog gene to provide a heterologous transcription regulator linked to the coding region of this gene. Administration of the ptc therapeutic agent is used to protect cerebral tissues against ischaemic injury; to treat cerebral infarct or ischaemia, stroke (thrombotic or embolic) and transient ischaemic attacks. It may also be used as a prophylactic in many other cases of injury to the brain or spinal cord, oedema caused by trauma, haemorrhage and encephalomyelitis, or in conjunction with (coronary bypass) surgery. Treatment (which may be prophylactic) is used where ischaemic/epoxic conditions may cause cerebral hypoxia, or progressive loss of neurons due to oxygen depletion, including in patients with hypotension. The treatment reduces CIV by at least 25, particularly at least 70%. The present sequence represents a hedgehog sequence given in the present invention

Query Match 83.6%; Score 2065; DB 2; Length 437;  
 Best Local Similarity 87.2%; Pred. No. 1.6e-180;  
 Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
 QY 1 MLLARCLLLVSLVCSGLACGPGFGKRRHPKLLTPLAYKQFINVAEKTIGASG 60  
 Db 2 LLLARCLFLVILASLLVCPGLACGPGFGKRRHPKLLTPLAYKQFINVAEKTIGASG 61  
 QY 61 RYEGKISRNSRPFKELTNNPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Db 62 RYEGKITRNSRPFKELTNNPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 121  
 QY 121 KLRVTEGDEDDHSESLHYEGRAVDITTSDRSKYGMRLARLAVAGFDWVYVESKAH 180  
 Db 122 RLRVTEGDEDDHSESLHYEGRAVDITTSDRSKYGMRLARLAVAGFDWVYVESKAH 181  
 QY 181 IHCSVKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGLLYSDFLT 240  
 Db 182 IHCSVKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGLLYSDFLT 241  
 QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
 Db 242 FLDRDEGAKKVFYVETREPRERLLITAAHLLFVAPHND-----SGTTPG 286  
 QY 301 PRALFASVRPGQRYVVAEDGDRLLPAAVHVTLSSEAGAYAPLTAQGTLLINRVL 360  
 Db 287 PSALFASVRPGQRYVVAEDGDRLLPAAVHVTLSSEAGAYAPLTAHGTLLINRVL 346  
 QY 361 ASCYAVIEHSHWAHRAFPFLAHALIAALAPARTDGGSGGDRGGGGRVALTAPGA 420  
 Db 347 ASCYAVIEHSHWAHRAFPFLAHALIAALAPARTD-----GGGGGSIIP-AAQSA 395  
 QY 421 ADAPGAGATAGHWSQLLYOIGTWLLDSEALHPLGMVAKSS 462  
 Db 396 TEARGAETAGHWSQLLYHIGTWLLDSEYHPLGMVAKSS 437

RESULT 37  
 ID AA05857  
 XX AA05857 standard; protein; 437 AA.  
 AC AA05857;  
 XX  
 DT 02-AUG-1999 (first entry)  
 DE Mouse sonic hedgehog Shh protein.  
 XX  
 KW Sonic hedgehog; Shh; mouse; epithelial tissue; epithelium;  
 KW cutaneous tissue; skin; hair; wound healing; vulnery; burn;  
 KW skin grafting; pressure sore; ulcer; ulcerative colitis; alopecia;  
 KW psoriasis; keratosis; acne; comedogenic lesion; folliculitis;  
 KW pseudofolliculitis; keratoacanthoma; callosities; Darier's disease; scar;  
 KW autoimmune disease; pemphigus; epidermolysis; lupus lesion;  
 KW desquamative lesion; carcinoma; therapy; hedgehog therapeutic;

ptc therapeutic; patched.  
 OS Mus musculus.  
 XX WO9920298-A1.  
 XX 29-APR-1999.  
 XX 20-OCT-1999; 98WO-US022227.  
 XX 20-OCT-1997; 97US-00955552.  
 PR 11-SEP-1999; 98US-00151999.  
 XX (ONTO-) ONTOGENY INC.  
 PA Wang EA;  
 PI WPI: 1999-288170/24.  
 DR N-PSDB; AAX25620.  
 XX Use of hedgehog polypeptides on patched therapeutics.  
 PS Claim 26; Page 124-125; 146pp; English.  
 CC The present sequence represents mouse Sonic hedgehog protein Shh. The invention relates to a method for modulating the growth state an epithelial cell by ectopically contacting the epithelial cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a hedgehog polypeptide or gene therapy construct) or ptc therapeutic (i.e. a small organic molecule that mimics the effect of hedgehog proteins on patched signalling, or activates or potentiates patched signalling) in an amount effective to alter the rate of proliferation of the epithelial cell. The hedgehog therapeutic preferably comprises at least a bioactive extracellular portion of a hedgehog protein (see AAX05854-62) encoded by a vertebrate hedgehog gene (see AAX25617-25), especially a human hedgehog gene.  
 CC Promotion of proliferation of epithelial cells can be used to control a wound healing process in e.g. burn treatment, skin regeneration, skin grafting, pressure sore treatment, dermal ulcer treatment, post surgery scar reduction or treatment of ulcerative colitis (claimed). It can also be used to induce hair growth for the treatment of alopecia (claimed).  
 CC Inhibition of the growth of epithelial tissue can be used to treat or prevent hyperplastic or neoplastic conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions, folliculitis and pseudofolliculitis, keratoacanthoma, callosities, Darier's disease, keloids, hypertrophic scars, or autoimmune disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus foliaceus, pemphigus vegetans, pemphigus erythematosus, epidermolysis, lupus lesions, desquamative lesions or carcinomas. The methods can also be used to counteract the effects of ageing on skin

Query Match 83.6%; Score 2065; DB 2; Length 437;  
 Best Local Similarity 87.2%; Pred. No. 1.6e-180;  
 Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
 QY 1 MLLARCLLLVSLVCSGLACGPGFGKRRHPKLLTPLAYKQFINVAEKTIGASG 60  
 Db 2 LLLARCLFLVILASLLVCPGLACGPGFGKRRHPKLLTPLAYKQFINVAEKTIGASG 61  
 QY 61 RYEGKISRNSRPFKELTNNPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
 Db 62 RYEGKITRNSRPFKELTNNPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 121  
 QY 121 KLRVTEGDEDDHSESLHYEGRAVDITTSDRSKYGMRLARLAVAGFDWVYVESKAH 180  
 Db 122 RLRVTEGDEDDHSESLHYEGRAVDITTSDRSKYGMRLARLAVAGFDWVYVESKAH 181  
 QY 181 IHCSVKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGLLYSDFLT 240  
 Db 182 IHCSVKAENSVAKSGGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGLLYSDFLT 241  
 QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300

Db 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASRVPRGQVYVVAERDGRLLPAVHVSVTLSAEAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQVYVVAERDGRLLPAVHVSVTLSAEAGAYAPLTAQGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTDGGSGDGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWSYQLLYOIGTWLSDSEALHPLGMAYKSS 462  
Db 396 TEARGAPTAGIHWSYQLLYOIGTWLSDSEALHPLGMAYKSS 437  
RESULT 39  
AAW97768  
ID AAW97768 standard; protein; 437 AA.  
XX AC AAW97768;  
XX DT 21-MAY-1999 (first entry)  
XX DE Mouse Sonic hedgehog (Shh) protein.  
XX KW Sonic hedgehog; Shh protein; mouse; dopaminergic; GABA-nergic;  
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;  
KW Huntington's disease; amyotrophic lateral sclerosis; cerebral ischaemia;  
KW hypoxia; neuroprotective; therapy.  
XX OS Mus sp.  
XX PN WO9904775-A2.  
XX PD 04-FEB-1999.  
XX PF 24-JUL-1998; 98WO-US015419.  
XX PR 24-JUL-1997; 97US-00900220.  
XX PA (ONTO-) ONTOGENY INC.  
XX PI Miao N, Wang M, Mahanthappa NK, Pang K;  
XX WPI: 1999-142578/12.  
XX DR N-PSDB; AAX07274.  
XX PT Increasing the survival of neuronal, dopaminergic and GABA-nergic cells -  
PT by using a ptc therapeutic such as a protein kinase inhibitor, or an  
PT agent derived from hedgehog polypeptides, useful in the treatment of  
PT Parkinson's disease.  
XX PS Disclosure; Page 89-91; 138pp; English.  
XX CC This polypeptide is mouse Shh Sonic hedgehog protein. The invention is  
CC based on the finding that hedgehog proteins are useful as protective  
CC agents in the treatment and prophylaxis of neurodegenerative disorders  
CC resulting from the loss of dopaminergic and/or GABA-nergic neurons, or  
CC the general loss of tissue from the substantia nigra. Exemplary disorders  
CC include Parkinson's disease, Huntington's disease (both claimed),  
CC ankyrotropic lateral sclerosis and cerebral ischaemia. The invention  
CC relates to hedgehog therapeutics (i.e. hedgehog polypeptides and gene  
CC therapy constructs e.g. constructs encoding recombinant hedgehog  
CC polypeptides and trans-activation constructs for altering hedgehog gene  
CC regulatory sequences) and ptc therapeutics (i.e. agents which mimic the  
CC effect of naturally occurring hedgehog proteins on patched signalling)  
CC that are effective in both human and animal subjects, Human Ikh and Dhh  
CC polypeptides (see AAW97763-64) are preferred. The products can also be  
CC used for the maintenance of differentiated neurons in cultures, and to  
CC enhance the implantation of such neuronal cells in an animal. They can be  
CC used to prevent or treat neurodegenerative conditions arising from the  
CC use of certain drugs, and in the prevention and/or treatment of hypoxia,  
CC e.g. as a neuroprotective agent

XX SQ Sequence 437 AA;  
Query March 83.68; Score 2065; DB 2; Length 437;  
Best Local Similarity 87.2%; Pred. No. 1.6e-180; Indels 26; Gaps 3;  
Matches 403; Conservative 9; Mismatches 24;  
Qy 1 MLLIARCLLLVLSVSLVCSGLACGPGRGFKRRHPKLTFLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLIARCLVILASSLLVCSGLACGPGRGFKRRHPKLTFLAYKQFIPNVAEKTILGASG 61  
Qy 61 RYEGKISNSERFKELTPNYPDIIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPDIIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 121  
Qy 121 KLRVTEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 122 RLRVTEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
Qy 181 IHCSVKAENSVAKSGCGPFGSATVHLEOGGTLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAKSGCGPFGSATVHLEOGGTLVKDLSFGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
Db 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASRVPRGQVYVVAERDGRLLPAVHVSVTLSAEAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQVYVVAERDGRLLPAVHVSVTLSAEAGAYAPLTAQGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTDGGSGDGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWSYQLLYOIGTWLSDSEALHPLGMAYKSS 462  
Db 396 TEARGAPTAGIHWSYQLLYOIGTWLSDSEALHPLGMAYKSS 437  
RESULT 39  
AAW97768  
ID AAW97768 standard; protein; 437 AA.  
XX AC AAW97768;  
XX DT 05-JUL-1999 (first entry)  
XX DE Mouse Sonic hedgehog protein Shh.  
XX KW Sonic hedgehog; Shh protein; mouse; hedgehog therapeutic;  
KW ptc therapeutic; patched; signal transduction; muscle atrophy; cachexia;  
KW muscular myopathy; myoblastic sarcoma; therapy.  
XX OS Mus sp.  
XX PN WO9910004-A2.  
XX PD 04-MAR-1999.  
XX PF 28-AUG-1998; 98WO-US017922.  
XX PR 29-AUG-1997; 97US-0057394P.  
XX PA (ONTO-) ONTOGENY INC.  
XX PI Bladen CS, Currie PD, Ingham PW, Hughes SM;  
XX WPI: 1999-243557/20.  
XX DR N-PSDB; AAX25101.  
XX PT A new method to regulate muscle growth.

PS Disclosure; Page 115-116; 130pp; English.

XX The present sequence is mouse Sonic hedgehog protein Shh. The invention  
CC relates to a method for modulating the formation and/or maintenance of  
CC muscle tissue by ecotopically contacting muscle cells, especially muscle  
CC stem/progenitor cells, in vitro or in vivo, with a hedgehog therapeutic  
CC (i.e. hedgehog polypeptides and gene therapy constructs) or ptc  
CC therapeutic (i.e. a small organic molecule that mimics the effect of  
CC hedgehog proteins on patched signalling, or activates or potentiates  
CC patched signalling) in an amount effective to alter the growth state of  
CC the treated cells. Also claimed is a method for treatment or prevention  
CC of disorders of, or surgical or cosmetic repair of, such muscle tissues,  
CC by administering a hedgehog polypeptide or ptc therapeutic. The disorder  
CC may be muscle atrophy, in particular skeletal muscle atrophy or cardiac  
CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The  
CC hedgehog polypeptide or ptc therapeutic can inhibit growth of myoblastic-  
CC derived tissue to provide treatment of hyperplastic or neoplastic growth  
CC of muscle tissue such as in myoblastic sarcoma (also claimed). The  
CC hedgehog therapeutic preferably comprises at least a bioactive  
CC extracellular portion of a hedgehog protein (see AAY05510-19) encoded by  
CC a vertebrate hedgehog gene (see AAX25098-107), especially a human  
CC hedgehog gene

XX Sequence 437 AA;

Query Match 83.6%; Score 2065; DB 2; Length 437;

Best Local Similarity 87.2%; Pred. No. 1.6e-180; Indels 26; Gaps 3;  
Matches 403; Conservative 9; Mismatches 24;

QY 1 MLLARCLLVLSVLLVCSGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTLCGASG 60  
DB 2 LLLARCFVLIVASSLLVCPGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTLCGASG 61  
QY 61 RYEGKISRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQPGV 120  
DB 62 RYEGKITRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQPGV 121  
QY 121 KLRVTEGDEQGHSESLHVEGRAVDITTDGDRSKYGMRLARLAVAGFDWVYVESKAH 180  
DB 122 KLRVTEGDEQGHSESLHVEGRAVDITTDGDRSKYGMRLARLAVAGFDWVYVESKAH 181  
QY 181 THCSVKAENSVAAGSGGFPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 182 THCSVKAENSVAAGSGGFPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETETREPRLILTAHLLFVAFHNDSTGTPRASSGGSPSGGALG 300  
DB 242 FLDRDEGAKKVFYVETETREPRLILTAHLLFVAFHNDSTGTPRASSGGSPSGGALG 286  
QY 301 PRALFASRVPRQRYVVAERDGRRLPAAVHVSVTLSEAGAYAPLTAQGTILINRVL 360  
DB 287 PSALFASRVPRQRYVVAERDGRRLPAAVHVSVTLSEAGAYAPLTAQGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHFAFPLAHALLAALAPARTDGGSDGGGGRGGRVALTAPGA 420  
DB 347 ASCYAVIEHSHWAHFAFPLAHALLAALAPARTDGGSDGGGGRGGRVALTAPGA 395  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLIDSEALHPLGMAVKSS 462  
DB 396 TEARGAETAGIHWYSQLLYHIGTWLIDSEALHPLGMAVKSS 437

RESULT 40

AAY96246

ID AAY96246 standard; protein; 437 AA.

XX AAY96246;

XX 11-SEP-2000 (first entry)

XX Partial mouse Shh.

XX Mouse; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;

KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;  
KW chronic inflammatory demyelinating polyneuropathy; CIDP; gene therapy;  
KW infection; inflammation; hereditary neuropathy;  
KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;  
KW multiple myeloma; nutritional imbalance; kidney disease;  
KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;  
KW Fabry's disease; Krabbe's disease; Metachromatic leukodystrophy;  
KW Tanager disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;  
KW Amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;  
KW Hereditary sensory neuropathy Type II; HSN II; B-cell lymphoma;  
KW Waldenstrom's Macroglobulinaemia; Chronic Lymphocytic Leukaemia;  
KW neuroprotective; cytoprotective; patched-mediated signal transduction.

XX Mus sp.

XX WO200027422-A2.

XX 18-MAY-2000.

XX 08-NOV-1999; 99WO-US026334.

XX 06-NOV-1998; 98US-00187387.

XX (BIOJ ) BIOGEN INC.

XX (ONTO-) ONTOGENY INC.

XX Galdes A, Mahanthappa N;

XX WPI; 2000-387341/33.

XX N-PSDB; AAA30277.

XX Novel method of preventing deterioration of peripheral nerves, useful for

XX treating or preventing neuropathy, e.g. where associated with diabetes or

XX viral infection, by administering hedgehog or patched agent.

XX Claim 7; Page 131-132; 152pp; English.

XX The present sequence is the partial mouse sonic hedgehog protein, Shh.  
CC This sequence inhibits expression of the patched gene which has been  
CC implicated in neuromuscular disorders (neuropathies). This sequence may  
CC therefore be used for treating neuromuscular disorders i.e. preventing  
CC degradation in function of motor or sensory nerves and protecting  
CC peripheral nerve cells under conditions that normally cause neuropathy. A  
CC variety of neuromuscular disorders may be treated: Guillain-Barre  
CC syndrome, GBS; peripheral neuropathy; diabetic neuropathy; alcohol-  
CC induced neuropathy; chronic inflammatory demyelinating polyneuropathy,  
CC CIDP; infection-induced neuropathy, including HIV infection; inflammation  
CC induced neuropathy; hereditary neuropathy e.g. Charcot-Marie-Tooth  
CC disease (CMT), Familial Amyloidotic neuropathy, Refsum's disease,  
CC Abetalipoproteinemia, Tangier disease, Krabbe's disease, Metachromatic  
CC leukodystrophy, Fabry's disease, Dejerine-Sottas syndrome, Hereditary  
CC sensory neuropathy Type II (HSN II) and Amyotrophic lateral sclerosis  
CC (ALS); acute neuropathy e.g. Miller-Fisher syndrome; neuropathy caused by  
CC vasculitis; neuropathy associated with tumours e.g. lung cancer, multiple  
CC myeloma, B-cell lymphoma, Waldenstrom's Macroglobulinaemia, Chronic  
CC Lymphocytic leukaemia; neuropathy associated with: amyloidosis,  
CC nutritional imbalance, kidney disease, trauma; and hypothyroid  
CC disorders

XX The coding sequence may be used in gene therapy of the above

XX SQ Sequence 437 AA;

Query Match 83.6%; Score 2065; DB 3; Length 437;

Best Local Similarity 87.2%; Pred. No. 1.6e-180; Indels 26; Gaps 3;

Matches 403; Conservative 9; Mismatches 24;

QY 1 MLLARCLLVLSVLLVCSGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTLCGASG 60

DB 2 LLLARCFVLIVASSLLVCPGLACGPGRGFKRHPKLTPLAYKQIPNVAEKTLCGASG 61

QY 61 RYEGKISRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQPGV 120

DB 62 RYEGKITRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQPGV 121

CC involves a direct or indirect antagonism of patched-mediated regulation  
CC of gene expression. This method is useful for the treatment or prevention  
CC of lung diseases, like cancer, cystic fibrosis, bronchopneumocystosis,  
CC bronchitis, bronchospasm, sarcoidosis, silicosis, eosinophilic granuloma,  
CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress  
CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary  
CC pulmonary hypertension. It is also used to control wound healing or other  
CC reformation processes in the lung and augment lung transplantation. The  
CC present sequence is the mouse sonic hedgehog (Shh) protein, essential for  
CC development of the respiratory system. Hedgehog polypeptides can be used  
CC to control the formation and/or maintenance of the lung tissue  
XX  
SQ

Query Match 83.8%; Score 2065; DB 3; Length 437;  
Best Local Similarity 87.2%; Pred. No. 1.6e-180;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLILARCLLLVSSLLVCSGLACGPGRGFKGRHKKLTPLAYKQFIPNVAEKTILGASG 60  
DB 2 LLLARCFVLASSLLVCSGLACGPGRGFKGRHKKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERFKELTPNYPNPIIFKDEENTGADRLMTORCKOKLNAIAISVMNQPGV 120  
DB 62 RYEGKITRNSERFKELTPNYPNPIIFKDEENTGADRLMTORCKOKLNAIAISVMNQPGV 121  
QY 121 KLRTGWDGDDGHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 180  
DB 122 LRLRTGWDGDDGHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 181  
QY 181 IHCSVKAENSVAKSGCGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAKSGCGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFVVIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPGGGALG 300  
DB 242 FLDRDEGAKKVFVVIETREPRERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQVVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAQTILINRVL 360  
DB 287 PSALFASRVPRGQVVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAHTGILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDGGSDGGDRGGGRVALTAPGA 420  
DB 347 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
RESULT 42  
AAAY70679 standard; protein; 437 AA.  
ID AAY70679  
XX  
AC AAY70679;  
XX  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Mouse Sonic hedgehog (Shh) protein.  
XX  
KW Sonic hedgehog; Shh; mouse; growth modulator; therapeutic agent; lung;  
KW hedgehog; hh; patched; ptc; fibroblast growth factor; fgf-10;  
KW antiproliferative; anticancer; vulvar; antirheumatic; hypotensive;  
KW anti-inflammatory; antiarthritic; tuberculostatic; asthma;  
KW antimicrobial; antiallergy; treatment; prevention; lung diseases; cancer;  
KW cystic fibrosis; bronchitis; emphysema; respiratory distress syndrome;  
KW tuberculosis; wound healing; lung transplantation.  
XX  
OS Mus sp.  
XX  
XX WO200015246-A2.  
XX  
XX 23-MAR-2000.  
XX  
XX 10-SEP-1999; 99WO-US020500.  
XX  
XX 11-SEP-1998; 98US-0099952P.  
XX  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Pepicelli C, Lewis P, McMahon AP;  
XX  
XX WPI; 2000-271252/23.  
XX  
XX N-PSDB; AAZ52260.  
XX  
XX Modulation of lung tissue or cell growth rate used for treating or  
XX preventing damage to lung tissue comprises ectopically contacting tissue  
XX with hedgehog therapeutic, patched therapeutic or fibroblast growth  
XX factor-10.  
XX  
XX Claim 14; Page 119-121; 143pp; English.  
XX  
XX The patent discloses a method for modulating the growth state of  
XX epithelial or mesenchymal cells of the lung, by ectopically contacting  
XX the tissue with a therapeutic agent, that can effectively alter the rate  
XX of proliferation of cells. This agent can be selected from hedgehog (hh),  
XX patched (ptc) or fibroblast growth factor (fgf)-10 therapeutics. It

CC involves a direct or indirect antagonism of patched-mediated regulation  
CC of gene expression. This method is useful for the treatment or prevention  
CC of lung diseases, like cancer, cystic fibrosis, bronchopneumocystosis,  
CC bronchitis, bronchospasm, sarcoidosis, silicosis, eosinophilic granuloma,  
CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress  
CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary  
CC pulmonary hypertension. It is also used to control wound healing or other  
CC reformation processes in the lung and augment lung transplantation. The  
CC present sequence is the mouse sonic hedgehog (Shh) protein, essential for  
CC development of the respiratory system. Hedgehog polypeptides can be used  
CC to control the formation and/or maintenance of the lung tissue  
XX  
SQ

Query Match 83.8%; Score 2065; DB 3; Length 437;  
Best Local Similarity 87.2%; Pred. No. 1.6e-180;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLILARCLLLVSSLLVCSGLACGPGRGFKGRHKKLTPLAYKQFIPNVAEKTILGASG 60  
DB 2 LLLARCFVLASSLLVCSGLACGPGRGFKGRHKKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERFKELTPNYPNPIIFKDEENTGADRLMTORCKOKLNAIAISVMNQPGV 120  
DB 62 RYEGKITRNSERFKELTPNYPNPIIFKDEENTGADRLMTORCKOKLNAIAISVMNQPGV 121  
QY 121 KLRTGWDGDDGHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 180  
DB 122 LRLRTGWDGDDGHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAH 181  
QY 181 IHCSVKAENSVAKSGCGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAKSGCGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFVVIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPGGGALG 300  
DB 242 FLDRDEGAKKVFVVIETREPRERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQVVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAQTILINRVL 360  
DB 287 PSALFASRVPRGQVVVAERDGRLLPAAVHSVTLSSEAAAGAYAPLTAHTGILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDGGSDGGDRGGGRVALTAPGA 420  
DB 347 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
RESULT 42  
AAAY95284 standard; protein; 437 AA.  
ID AAY95284  
XX  
AC AAY95284;  
XX  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE Mouse Sonic hedgehog Shh protein.  
XX  
KW Sonic hedgehog; Shh; mouse; excitotoxicity; Parkinson's disease;  
KW Huntington's disease; neuronal degeneration; neuroprotective;  
KW dopaminergic; GABAergic; substantia nigra; therapy.  
XX  
OS Mus musculus.  
XX  
XX WO200035948-A1.  
XX  
XX 22-JUN-2000.  
XX  
XX 03-DEC-1999; 99WO-US028721.  
XX



CC deliver polynucleotides encoding these lipid modulators

XX Sequence 437 AA;

SQ Query Match 83.6%; Score 2065; DB 3; Length 437;  
Best Local Similarity 87.2%; Pred. No. 1.6e-180;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

Qy 1 MLLIARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLFVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
Qy 61 RYEGKISNSRERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
Db 62 RYEGKITNSRERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 121  
Qy 121 KLRTVEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 122 RLAVTEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
Qy 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGKTLVKDLSGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGKTLVKDLSGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDGAKKVYVIETREPRERLLITAAHLLFVAPHND-----SGPTPG 300  
Db 242 FLDRDGAKKVYVIETREPRERLLITAAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASVRPQGVYVVAERDGRRLPAAVHVSITLSEEAAGAYAPLTAQTILINRVL 360  
Db 287 PSALFASVRPQGVYVVAERDGRRLPAAVHVSITLSEEAAGAYAPLTAQTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGSDGSGGGRGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWSQLLYOIGTWLDDSEALHPLGMVAVKSS 462  
Db 396 TEARGAEPGATAGIHWSQLLYOIGTWLDDSEALHPLGMVAVKSS 437

RESULT 44

AAB84672

ID AAB84672 standard; protein; 437 AA.

XX AC AAB84672;

XX DT 17-SEP-2001 (first entry)

XX DE Amino acid sequence of a mouse hedgehog (Shh) polypeptide.

XX Shh gene; hedgehog gene; T lymphocyte; patched gene; infection; diabetes;  
KW nutritional deficiency; graft rejection; hyperacute response;  
KW cornea transplant; autoimmune disorder; multiple sclerosis; psoriasis;  
KW atopic dermatitis; inflammatory disease; proliferative disease;  
KW hyperproliferative disease; eczematous dermatitis; urticaria; vasculitis;  
KW scleroderma.

OS Mus sp.

XX WO200140438-A2.

XX PN 07-JUN-2001.

XX PD 30-NOV-2000; 2000WO-US032590.

XX PF 30-NOV-1999; 99US-0168112P.

XX PR (CURI-) CURIS INC.

XX XX Crompton T;

XX WPI; 2001-441484/47.

DR N-PSDB; AAB28449.

XX Modulating immune function comprises administration of a hedgehog or ptc  
PT agent, for treating e.g. diabetes, eczematous dermatitis, urticaria or  
PT vasculitis.

XX Claim 4; Page 81-82; 105pp; English.

XX The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog  
CC gene products and signal transduction pathways involving hedgehog are  
CC involved in the maturation of T lymphocytes. The specification describes  
CC a method for modulating immune function, by administration of a hedgehog  
CC or patched (ptc) polypeptide, agonists or antagonists. The method is used  
CC to treat disorders affecting the regulation of lymphocytes, particularly  
CC maturation and/or activation of T lymphocytes. It is used to treat  
CC bacterial or viral infection, diabetes, nutritional deficiencies, graft  
CC rejection or other hyperacute response such as kidney, heart, lung, bone  
CC marrow spleen skin or cornea transplant or autoimmune disorders such as  
CC multiple sclerosis, psoriasis or atopic dermatitis. The method is used to  
CC treat inflammatory, proliferative and hyperproliferative diseases, as  
CC well as cutaneous manifestations of immunological disorders such as  
CC eczematous dermatitis, urticaria, vasculitis and scleroderma

XX Sequence 437 AA;

SQ Query Match 83.6%; Score 2065; DB 4; Length 437;

Best Local Similarity 87.2%; Pred. No. 1.6e-180;

Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

Qy 1 MLLIARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60

Db 2 LLLARCLFVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61

Qy 61 RYEGKISNSRERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120

Db 62 RYEGKITNSRERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 121

Qy 121 KLRTVEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180

Db 122 RLRTVEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181

Qy 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGKTLVKDLSGDRVLAADDQGRLLYSDFLT 240

Db 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGKTLVKDLSGDRVLAADDQGRLLYSDFLT 241

Qy 241 FLDRDGAKKVYVIETREPRERLLITAAHLLFVAPHND-----SGPTPG 300

Db 242 FLDRDGAKKVYVIETREPRERLLITAAHLLFVAPHND-----SGPTPG 286

Qy 301 PRALFASVRPQGVYVVAERDGRRLPAAVHVSITLSEEAAGAYAPLTAQTILINRVL 360

Db 287 PSALFASVRPQGVYVVAERDGRRLPAAVHVSITLSEEAAGAYAPLTAQTILINRVL 346

Qy 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGSDGSGGGRGGRVALTAPGA 420

Db 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395

Qy 421 ADAPGAGATAGIHWSQLLYOIGTWLDDSEALHPLGMVAVKSS 462

Db 396 TEARGAEPGATAGIHWSQLLYOIGTWLDDSEALHPLGMVAVKSS 437

RESULT 45

AAB60263

ID AAB60263 standard; protein; 437 AA.

XX AC AAB60263;

XX DT 30-MAR-2001 (first entry)

XX DE Mouse Sonic hedgehog (Shh) protein, SEQ ID NO:13.

XX Hedgehog protein; polymer conjugate; polyalkene glycol group;





Db 182 IHCSVKAENSVAAGSGGPGSATVHLEQGGTKLVKDLRPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERILLTAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPGQRYVVAERDGRRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPGQRYVVAERGDRRLLPAAVHVSVTLSREBEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDRGSDGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGIHWSQLLYHIGTWLDDSETMHPGLMAVKSS 437  
RESULT 48  
AAB85736  
ID AAB85736 standard; protein; 437 AA.  
AC AAB85736;  
XX 29-OCT-2001 (first entry)  
DT Mouse sonic hedgehog (Shh) polypeptide.  
DE Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh; Dhh;  
KW desert hedgehog; cell differentiation; mouse.  
XX Mus sp.  
OS US6271363-B1.  
FN 07-AUG-2001.  
PD 20-OCT-1997; 97US-00954698.  
PF 30-DEC-1993; 93US-00176427.  
PR 14-DEC-1994; 94US-00356060.  
PR 04-MAY-1995; 95US-00435093.  
PR 05-JUN-1995; 95US-00462386.  
XX (HARD ) HARVARD COLLEGE  
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX Ingham PW, McMahon AP, Tabin CJ;  
XX WPI; 2001-456723/49.  
DR N-PSDB; AAB76110.  
XX Novel nucleic acid encoding a hedgehog polypeptide, used to produce the  
PT polypeptide, which is used to promote proliferation, survival, and/or  
PT differentiation of neuronal and mesodermal tissue.  
XX Claim 1; Col 127-130; 118pp; English.  
XX The invention relates to nucleic acids encoding hedgehog proteins  
CC selected from sonic hedgehog (Shh), indian hedgehog (Ihh), desert  
CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the  
CC formation of ordered spatial arrangements of differentiated tissue in  
CC vertebrates. The nucleic acid sequences are useful for producing hedgehog  
CC proteins, used for promoting differentiation of, or survival of  
CC differentiated, neuronal cells, and for promoting proliferation, survival  
CC or differentiation of mesenchymal, endodermal or ectodermal tissue,  
CC particularly chondrocytes, or testicular germ line cells. The present  
XX sequence represents a mouse Shh polypeptide  
SQ Sequence 437 AA;

Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 1.6e-180; Indels 26; Gaps 3;  
Matches 403; Conservative 9; Mismatches 24;  
QY 1 MLLARCLLLVLSLLVCSGLACPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTGLGASG 60  
Db 2 LLLARCLFVILASSLLVCPGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTGLGASG 61  
QY 61 RYEGKISNSERFKELTNYNPDIIIFKDEENTGADRLMTQCKDKLNALAI SVMNQWPGV 120  
Db 62 RYEGKITRNSERFKELTNYNPDIIIFKDEENTGADRLMTQCKDKLNALAI SVMNQWPGV 121  
QY 121 KLURVTGDEDEGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
Db 122 KLURVTGDEDEGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGGPGSATVHLEQGGTKLVKDLRPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGGPGSATVHLEQGGTKLVKDLRPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERILLTAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPGQRYVVAERDGRRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPGQRYVVAERGDRRLLPAAVHVSVTLSREBEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDRGSDGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGIHWSQLLYHIGTWLDDSETMHPGLMAVKSS 437  
RESULT 49  
AAG65746  
ID AAG65746 standard; protein; 437 AA.  
AC AAG65746;  
XX 07-JAN-2002 (first entry)  
DT Mouse sonic hedgehog (Shh) polypeptide.  
DE Adipocyte; hedgehog polypeptide; desert hedgehog; indian hedgehog; Dhh;  
KW Ihh; sonic hedgehog; Shh; therapeutic; cytostatic; tumor.  
XX Mus sp.  
OS WO200164238-A2.  
FN 07-SEP-2001.  
PD 28-FEB-2001; 2001WO-US006450.  
PF 29-FEB-2000; 2000US-0186058P.  
PR (CURI-) CURIS INC.  
XX Zehentner B, Leser-Reiff U, Burtscher H;  
XX WPI; 2001-607352/69.  
DR N-PSDB; AAL66774.  
XX Method for regulating formation and/or maintenance of adipocyte tissue by  
PT contacting pre-adipocyte or adipocyte cells with a hedgehog polypeptide  
PT or ptc therapeutic.  
XX Disclosure; Page 97-99; 132pp; English.  
XX



Db 347 ASCVAVIEHSWAHRAFAFELAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWYSOLLYQIGTWLDDSEALHPLGNVAKSS 462  
Db 396 TEARGABFTAGIHWYSOLLYHIGTWLDDSETMHPGNVAKSS 437

Search completed: March 29, 2004, 18:32:03  
Job time : 64 secs



101 1234 50.0 411 4 US-09-639-695-10 Sequence 10, Appl  
102 1234 50.0 411 4 US-09-448-188-10 Sequence 10, Appl  
103 1234 50.0 411 4 US-08-954-128-10 Sequence 10, Appl  
104 1234 50.0 411 4 US-09-704-917-12 Sequence 12, Appl  
105 1234 50.0 411 4 US-08-954-740-10 Sequence 10, Appl  
106 1234 50.0 411 4 US-09-351-999-12 Sequence 12, Appl  
107 1234 50.0 411 4 US-09-736-476-10 Sequence 10, Appl  
108 1183 47.9 396 1 US-08-176-427B-4 Sequence 4, Appl  
109 1183 47.9 396 2 US-08-356-060A-9 Sequence 9, Appl  
110 1183 47.9 396 3 US-08-460-900C-9 Sequence 9, Appl  
111 1183 47.9 396 3 US-08-674-509B-9 Sequence 9, Appl  
112 1183 47.9 396 3 US-08-954-698-9 Sequence 9, Appl  
113 1183 47.9 396 4 US-08-323-505-13 Sequence 13, Appl  
114 1183 47.9 396 4 US-08-957-874-9 Sequence 9, Appl  
115 1183 47.9 396 4 US-09-325-256-18 Sequence 18, Appl  
116 1183 47.9 396 4 US-09-639-695-9 Sequence 9, Appl  
117 1183 47.9 396 4 US-09-448-188-9 Sequence 9, Appl  
118 1183 47.9 396 4 US-08-954-128-9 Sequence 9, Appl  
119 1183 47.9 396 4 US-09-704-917-11 Sequence 11, Appl  
120 1183 47.9 396 4 US-08-954-740-9 Sequence 9, Appl  
121 1183 47.9 396 4 US-08-151-999-11 Sequence 11, Appl  
122 1183 47.9 396 4 US-09-736-476-9 Sequence 9, Appl  
123 1177 47.7 396 4 US-09-325-256-24 Sequence 24, Appl  
124 1177 47.7 396 4 US-09-704-917-17 Sequence 17, Appl  
125 1177 47.7 396 4 US-09-151-999-17 Sequence 17, Appl  
126 1031 41.8 221 3 US-08-460-900C-40 Sequence 40, Appl  
127 1031 41.8 221 3 US-08-954-698-40 Sequence 40, Appl  
128 1031 41.8 221 4 US-08-957-874-40 Sequence 40, Appl  
129 1031 41.8 221 4 US-09-325-256-15 Sequence 15, Appl  
130 1031 41.8 221 4 US-09-639-695-40 Sequence 40, Appl  
131 1031 41.8 221 4 US-09-448-188-40 Sequence 40, Appl  
132 1031 41.8 221 4 US-08-954-128-40 Sequence 40, Appl  
133 1031 41.8 221 4 US-09-704-917-21 Sequence 21, Appl  
134 1031 41.8 221 4 US-08-954-740-40 Sequence 40, Appl  
135 1031 41.8 221 4 US-09-151-999-21 Sequence 21, Appl  
136 1031 41.8 221 4 US-09-736-476-40 Sequence 40, Appl  
137 1027.5 41.6 336 1 US-08-176-427B-6 Sequence 6, Appl  
138 1027.5 41.6 336 2 US-08-356-060A-10 Sequence 10, Appl  
139 1027.5 41.6 336 3 US-08-674-509B-10 Sequence 2, Appl  
140 932 37.7 174 4 US-09-325-256-2 Sequence 14, Appl  
141 930.5 37.7 312 3 US-08-460-900C-14 Sequence 14, Appl  
142 930.5 37.7 312 3 US-08-954-698-14 Sequence 14, Appl  
143 930.5 37.7 312 4 US-09-639-695-14 Sequence 14, Appl  
144 930.5 37.7 312 4 US-09-448-188-14 Sequence 14, Appl  
145 930.5 37.7 312 4 US-08-954-128-14 Sequence 14, Appl  
146 930.5 37.7 312 4 US-08-954-740-14 Sequence 14, Appl  
147 930.5 37.7 312 4 US-08-356-060A-14 Sequence 14, Appl  
148 930.5 37.7 313 4 US-08-957-874-14 Sequence 14, Appl  
149 930.5 37.7 313 4 US-09-736-476-14 Sequence 14, Appl  
150 930.5 37.7 313 4

## ALIGNMENTS

RESULT 1  
US-08-356-060A-13 Application US/08356060A  
; Sequence 13, Application US/08356060A  
; Patent No. 5844079  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,060A  
FILING DATE: 14-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMI-006CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-060A-13  
Query Match 99.9%; Score 2467; DB 2; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238; Indels 0; Gaps 0;  
Matches 475; Conservative 0; Mismatches 0;  
QY 1 MLLARCLLLVLSVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTLAGS 60  
Db 1 MLLARCLLLVLSVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTLAGS 60  
QY 61 RYEGKISNSRFRKLTPTNPNPDIIFKDEENTGADRLMTORCKDKNALAISVNNQPGV 120  
Db 61 RYEGKISNSRFRKLTPTNPNPDIIFKDEENTGADRLMTORCKDKNALAISVNNQPGV 120  
QY 121 KLRTVEGDEDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVAEAGFDWVYESKAH 180  
Db 121 KLRTVEGDEDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVAEAGFDWVYESKAH 180  
QY 181 IHCSYKAENSVAAGSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSYKAENSVAAGSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPSPGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPSPGGALG 300  
QY 301 PRALFASVRPGQRYVVAERDGRRLLPAAVHVSVTLSSEAGAYAPLTAQGTLLINRVL 360  
Db 301 PRALFASVRPGQRYVVAERDGRRLLPAAVHVSVTLSSEAGAYAPLTAQGTLLINRVL 360  
QY 361 ASCYAVIEHSHWAHPAFAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTPGA 420  
Db 361 ASCYAVIEHSHWAHPAFAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSXSRGACGGAREGA 475  
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSXSRGACGGAREGA 475  
RESULT 2  
US-08-460-900C-13 Application US/08460900C  
; Sequence 13, Application US/08460900C  
; Patent No. 6165747  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; APPLICANT: Bumcrot, David A.

APPLICANT: Marti-Gorostiza, Elisa  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related thereto  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/460,900C  
APPLICATION NUMBER: US/08/460,900C  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994

APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-900C-13

Query Match 99.9%; Score 2467; DB 3; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLLVLSVLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 1 MLLARCLLLVLSVLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60

Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120

Qy 121 KLRTVEGDWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 121 KLRTVEGDWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180

Qy 181 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

Qy 241 FLDRDDGAKKYFYVITETREPRERLLTAHLLFVAPHNDSATGPEASSGGPPSGGALG 300  
Db 241 FLDRDDGAKKYFYVITETREPRERLLTAHLLFVAPHNDSATGPEASSGGPPSGGALG 300

Qy 301 PRALPASVRVQRVYVVAERDGRRLLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALPASVRVQRVYVVAERDGRRLLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360

Qy 361 ASCVAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGRVALTAPGA 420

Db 361 ASCVAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGRVALTAPGA 420

Qy 421 ADAPGATAGIHWYSOLLVOIGTWLDDSEALHPLGNVAVKSSXSRGAGGAREGA 475

Db 421 ADAPGATAGIHWYSOLLVOIGTWLDDSEALHPLGNVAVKSSXSRGAGGAREGA 475

RESULT 3  
US-08-674-509B-13  
Sequence 13, Application US/08674509B  
Patent No. 6261786  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
APPLICANT: Marigo, Valeria  
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEHOG AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/674,509B  
APPLICATION NUMBER: US/08/674,509B  
FILING DATE: 02-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/460,900  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-509B-13

Query Match 99.9%; Score 2467; DB 3; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLARCLLLVLSVLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 1 MLLARCLLLVLSVLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60

Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120

Qy 121 KLRTVEGDWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 121 KLRTVEGDWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180

Qy 181 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

QY 241 FLRDDGAKKVFYVETREPRERLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
DB 241 FLRDDGAKKVFYVETREPRERLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASRVPRGQVYVVAERDGRLLPAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAPAPRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
DB 361 ASCYAVIEEHSWAHRAPAPRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVKSXSRSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVKSXSRSRGAGGAREGA 475

## RESULT 4

US-08-954-698-13  
; Sequence 13, Application US/08954698  
; Patent No. 6271363  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/954,698  
; FILING DATE: 20-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/462,386  
; FILING DATE: 05-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 04-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMV-006.10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-954-698-13

Query Match 99.9%; Score 2467; DB 3; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVYSSLLVCSGLACGPGRGFKRRHPKLTPLAYKOFIPNVAEKTLCASG 60  
DB 1 MLLARCLLLVYSSLLVCSGLACGPGRGFKRRHPKLTPLAYKOFIPNVAEKTLCASG 60  
QY 61 RYEGKISRNSERFKELTPNYPDIIIPKDEBNTGADRLMTORCKDKLNALAISSVMNQPGV 120  
DB 61 RYEGKISRNSERFKELTPNYPDIIIPKDEBNTGADRLMTORCKDKLNALAISSVMNQPGV 120  
QY 121 KLRYTEGWDGDSHSESLHYEGRVDTITSDRRDSKYGMRLAVEAGFDWVYYSKAH 180  
DB 121 KLRYTEGWDGDSHSESLHYEGRVDTITSDRRDSKYGMRLAVEAGFDWVYYSKAH 180  
QY 181 IHCSVKAENSVAKSGCGCFPGSATVHLEQGTKLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGCGCFPGSATVHLEQGTKLVKDLSFGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLRDDGAKKVFYVETREPRERLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
DB 241 FLRDDGAKKVFYVETREPRERLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASRVPRGQVYVVAERDGRLLPAVHVSRTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAPAPRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
DB 361 ASCYAVIEEHSWAHRAPAPRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVKSXSRSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVKSXSRSRGAGGAREGA 475

RESULT 5  
US-08-957-874-13  
; Sequence 13, Application US/08957874  
; Patent No. 6384192  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/957,874  
; FILING DATE: 20-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/462,386  
; FILING DATE: 5-JUNE-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 4-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMV-006.10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.09  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-957-874-13

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
DB 1 MLLIARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCOKKLNALAI SVMNQWPGV 120  
DB 61 RYEGKISRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCOKKLNALAI SVMNQWPGV 120  
QY 121 KLRVTEGDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
DB 121 KLRVTEGDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
QY 181 IHCSVKAENSVAAGSGGCPGSAATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAGSGGCPGSAATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSSGSPSGGALG 300  
DB 241 FLDRDDGAKKVFYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSSGSPSGGALG 300  
QY 301 PRALFASVRPQGVVVAERDGRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASVRPQGVVVAERDGRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420  
DB 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSOLLYQIGTWLIDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSOLLYQIGTWLIDSEALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 6  
US-09-325-256-22  
Sequence 22, Application US/09325256  
Patent No. 6444793  
GENERAL INFORMATION:  
APPLICANT: PEPINSKY, R. BLAKE  
APPLICANT: BAKER, DARREN P.  
APPLICANT: WEN, DINGYI  
APPLICANT: WILLIAMS, KEVIN P.  
APPLICANT: GARGER, ELLEN A.  
APPLICANT: TAYLOR, FREDERICK R.  
APPLICANT: CALDES, ALPHONSE  
APPLICANT: PORTER, JEFFREY  
TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND METHODS  
FILE REFERENCE: BIV-067.01  
CURRENT APPLICATION NUMBER: US/09/325,256  
CURRENT FILING DATE: 1999-06-03  
PRIOR APPLICATION NUMBER: 60/099,800  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/078,935  
PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/089,695  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/067,423  
PRIOR FILING DATE: 1997-12-03  
PRIOR APPLICATION NUMBER: PCT/US98/25676  
PRIOR FILING DATE: 1998-12-03  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 22  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Xaa may be any amino acid  
US-09-325-256-22

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
DB 1 MLLIARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCOKKLNALAI SVMNQWPGV 120  
DB 61 RYEGKISRNSRERFKELTPNYPDIIFKDEENTGADRLMTQRCOKKLNALAI SVMNQWPGV 120  
QY 121 KLRVTEGDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
DB 121 KLRVTEGDEDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
QY 181 IHCSVKAENSVAAGSGGCPGSAATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAGSGGCPGSAATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSSGSPSGGALG 300  
DB 241 FLDRDDGAKKVFYIETREPRERLLTAAHLLFVAPHNDSATGEPEASSSGSPSGGALG 300  
QY 301 PRALFASVRPQGVVVAERDGRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASVRPQGVVVAERDGRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420  
DB 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSOLLYQIGTWLIDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSOLLYQIGTWLIDSEALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 7  
US-09-639-695-13  
Sequence 13, Application US/09639695  
Patent No. 6576237  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
McMahon, Andrew P.  
Tabin, Clifford J.  
Bumcrot, David A.  
Martí-Gorostiza, Elisa  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/639,695  
FILING DATE: 16-Aug-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.05  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-639-695-13

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSVSLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTILGASG 60  
DB 1 MLLARCLLLVSVSLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTILGASG 60  
QY 61 RYEGKISNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
DB 61 RYEGKISNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
QY 121 KLRVTEGWDEGHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKAH 180  
DB 121 KLRVTEGWDEGHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKAH 180  
QY 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYIETREPRERLLTAAHLFFVAPHNDSATGEPEASSGSGPPSGGALG 300  
DB 241 FLDRDDGAKKVFYIETREPRERLLTAAHLFFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPQGVVVAERDGRRLPPAAVHVTLSSEAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASVRPQGVVVAERDGRRLPPAAVHVTLSSEAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSWAHRAFPRLAHALLAALAPARTDRGSDSGDGGGRGGRVALTAPGA 420  
DB 361 ASCYAVIEHSWAHRAFPRLAHALLAALAPARTDRGSDSGDGGGRGGRVALTAPGA 420  
QY 421 ADAPGATAGIHWYSOLLVIGTGLWLDSEALPLGNVAVKSSRGAGGAREGA 475  
DB 421 ADAPGATAGIHWYSOLLVIGTGLWLDSEALPLGNVAVKSSRGAGGAREGA 475

RESULT 8  
US-09-448-188-13  
Sequence 13 Application US/09448188  
Patent No. 6607913  
GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.  
McMahon, Andrew P.  
Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/448,188  
FILING DATE: 23-No. 6607913-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-448-188-13

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.9e-238;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSVSLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTILGASG 60  
DB 1 MLLARCLLLVSVSLVCSGLACGPGRGKRRHPKLTPLAYKQFINVAEKTILGASG 60  
QY 61 RYEGKISNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
DB 61 RYEGKISNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120  
QY 121 KLRVTEGWDEGHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKAH 180  
DB 121 KLRVTEGWDEGHHSSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYESKAH 180  
QY 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYIETREPRERLLTAAHLFFVAPHNDSATGEPEASSGSGPPSGGALG 300  
DB 241 FLDRDDGAKKVFYIETREPRERLLTAAHLFFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPQGVVVAERDGRRLPPAAVHVTLSSEAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASVRPQGVVVAERDGRRLPPAAVHVTLSSEAGAYAPLTAQGTILINRVL 360

[illegible]

```

RESULT 9
US-08-954-128-13
; Sequence 13, Application US/08954128
; Patent No. 6610656
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170

```

```

Query Match      99.9%; Score 2467; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. NO. 1.9e-238;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLLIARCLLLVLLVSLVCSGLACGPGRGFKRRHPKKLTPLAYKQFIPNVAETLGASG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MLLIARCLLLVLLVSLVCSGLACGPGRGFKRRHPKKLTPLAYKQFIPNVAETLGASG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 RYEGKISGRNSERFKELTPNYPNDIIIFKDEENTGADRLMTQRCCKDLNALAISVNNQPGV 120

```

61	RYE	GKIS	RNS	RSF	KEL	TPN	YND	II	FKD	ENT	GAD	RLM	TQR	CKD	KL	NAL	AI	SV	MNQ	PGV	120																														
121	KUR	TEG	WDE	DGH	SES	LHY	EGR	AVD	IT	TS	DR	SKY	GM	LAR	LA	VE	AG	F	DV	WY	ESK	180																													
121	KUR	TEG	WDE	DGH	SES	LHY	EGR	AVD	IT	TS	DR	SKY	GM	LAR	LA	VE	AG	F	DV	WY	ESK	180																													
181	HC	S	VKA	NS	VAA	KSG	CG	FG	S	AT	VH	LE	OG	GT	K	V	DL	SP	GR	VLA	AD	QO	GR	LL	YSD	FL	240																								
181	HC	S	VKA	NS	VAA	KSG	CG	FG	S	AT	VH	LE	OG	GT	K	V	DL	SP	GR	VLA	AD	QO	GR	LL	YSD	FL	240																								
241	FL	DR	DG	G	A	K	K	Y	F	Y	I	E	T	R	P	R	L	L	T	A	A	H	L	L	F	V	A	PH	NS	D	S	A	T	C	E	P	P	A	S	S	G	S	G	F	P	P	G	G	A	L	300
241	FL	DR	DG	G	A	K	K	Y	F	Y	I	E	T	R	P	R	L	L	T	A	A	H	L	L	F	V	A	PH	NS	D	S	A	T	C	E	P	P	A	S	S	G	S	G	F	P	P	G	G	A	L	300
301	PR	AL	FA	S	R	V	R	P	Q	R	V	V	V	A	E	D	G	R	L	L	P	A	A	V	S	V	T	L	S	E	A	A	G	A	P	L	T	A	Q	T	I	L	I	N	R	V	L	360			
301	PR	AL	FA	S	R	V	R	P	Q	R	V	V	V	A	E	D	G	R	L	L	P	A	A	V	S	V	T	L	S	E	A	A	G	A	P	L	T	A	Q	T	I	L	I	N	R	V	L	360			
361	AS	C	Y	A	V	I	E	H	S	W	A	R	A	P	F	L	A	H	L	L	A	A	L	A	P	A	R	T	D	G	S	G	G	D	R	G	G	G	G	G	R	V	A	T	A	P	A	420			
361	AS	C	Y	A	V	I	E	H	S	W	A	R	A	P	F	L	A	H	L	L	A	A	L	A	P	A	R	T	D	G	S	G	G	D	R	G	G	G	G	G	R	V	A	T	A	P	A	420			
421	AD	AP	G	A	G	A	T	A	G	I	H	W	S	O	L	L	Y	O	I	G	T	W	L	D	S	E	A	L	P	L	G	M	V	K	S	X	S	R	G	A	G	G	A	B	E	G	A	475			
421	AD	AP	G	A	G	A	T	A	G	I	H	W	S	O	L	L	Y	O	I	G	T	W	L	D	S	E	A	L	P	L	G	M	V	K	S	X	S	R	G	A	G	G	A	B	E	G	A	475			

```

RESULT 10
US-09-704-917-15
; Sequence 15, Application US/09704917
; Patent No. 6616926
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Burkly, Linda
; APPLICANT: Wang, Li Chun
; TITLE OF INVENTION: METHODS OF MODULATING LIPID METABOLISM AND STORAGE
; FILE REFERENCE: A069PCT
; CURRENT APPLICATION NUMBER: US/09/704,917
; CURRENT FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/122,640
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 60/124,446
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn. Ver. 2.1
; SEQ ID NO 15
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid
US-09-704-917-15

```

Query Match	99.9%	Score 2467	DB 4	Length 475
Best Local Similarity	100.0%	Pred. No. 1.9e+238		
Matches 475	Conservative	0	Mismatches 0	Gaps 0
			Indels 0	

Qy	61	RYEGKISRNSERFKELTNNYPNDITTFDDEENTGADRLMTORCKDKLNAIAISVMNQWPGV	120
Db	61	RYEGKISRNSERFKELTNNYPNDITTFDDEENTGADRLMTORCKDKLNAIAISVMNQWPGV	120
Qy	121	KURVTGWDDEDHGHEESLHYEGRAVDITTSDDRSKYGMLAIAVEAGFDWYYESKAH	180
Db	121	KURVTGWDDEDHGHEESLHYEGRAVDITTSDDRSKYGMLAIAVEAGFDWYYESKAH	180
Qy	181	IHCVKAENSVAAKSGGCFPGSATVHLEEQGQTKLVKDLSPGDRVLAADQGRLLYSDFT	240
Db	181	IHCVKAENSVAAKSGGCFPGSATVHLEEQGQTKLVKDLSPGDRVLAADQGRLLYSDFT	240

[illegible]

RESULT 11  
US-08-954-740-13  
; Sequence 13, Application US/08954740  
; Patent No. 6630148  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

QY 181 IHCSVKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
DB 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSEEAAGAYAPLTAOGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYOIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSQLLYOIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475

## RESULT 13

US-09-736-476-13  
; Sequence 13, Application US/09736476  
; Patent No. 6664075  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; McMahon, Andrew P.  
; Tabin, Clifford J.  
; Bumcrot, David A.  
; Marti-Gorostiza, Elisa  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; Proteins and Uses Related Thereto

## NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/736,476  
FILING DATE: 13-Dec-2000  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMI-006CP4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO: 13:

## SEQUENCE CHARACTERISTICS:

LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-736-476-13

## Query Match 99.9%; Score 2467; DB 4; Length 475;

Best Local Similarity 100.0%; Pred. No. 1.9e-238; Mismatches 0; Indels 0; Gaps 0;  
Matches 475; Conservative 0;

QY 1 MLLARCLLLVLVSSLLVCSGLACGPGRGFGKRRHPKLLTPLAYKQFIPNVAEKTILGASG 60  
DB 1 MLLARCLLLVLVSSLLVCSGLACGPGRGFGKRRHPKLLTPLAYKQFIPNVAEKTILGASG 60  
QY 61 RYEGKISNSERFKELTNNYNPDIIIFKDEENTGADRLMTQCKOKLNALAISSVMNQWFGV 120  
DB 61 RYEGKISNSERFKELTNNYNPDIIIFKDEENTGADRLMTQCKOKLNALAISSVMNQWFGV 120  
QY 121 KLRTVTEGDEHGHSEESLHYEGRAVDITTSDRSKYGMRLARLAVEAGFDWVYYESKAH 180  
DB 121 KLRTVTEGDEHGHSEESLHYEGRAVDITTSDRSKYGMRLARLAVEAGFDWVYYESKAH 180  
QY 181 IHCSVKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
DB 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSEEAAGAYAPLTAOGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYOIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSQLLYOIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475

## RESULT 14

US-08-748-591-4  
; Sequence 4, Application US/08748591  
; Patent No. 5759811

## GENERAL INFORMATION:

APPLICANT: Epstein, Ervin  
APPLICANT: Hu, Zhilan  
APPLICANT: Bonifas, Jeanette  
TITLE OF INVENTION: Mutant Human Hedgehog Gene  
NUMBER OF SEQUENCES: 23

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish and Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748,591

## FILING DATE:

## CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: 06510/067001

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 322-5070

TELEFAX: (415) 854-0875

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 462 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-748-591-4

Query Match      97.5%; Score 2407; DB 1; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.9e-232;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60
QY 61 RYEGKISRNSERFKELTPNPNPDIIPKDEBNTGADRLMTORCKDKLNALAI SVMNQPGV 120
DB 61 RYEGKISRNSERFKELTPNPNPDIIPKDEBNTGADRLMTORCKDKLNALAI SVMNQPGV 120
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180
DB 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
QY 241 FLDRDDGAKKVFYVETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
DB 241 FLDRDDGAKKVFYVETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAHVSVTLSEEAAGAYAPLTAQGTILINRVL 360
DB 301 PRALFASRVPRGQRYVVAERDGRLLPAHVSVTLSEEAAGAYAPLTAQGTILINRVL 360
QY 361 ASCVAVIEEHSWAHRAFPAPFLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420
DB 361 ASCVAVIEEHSWAHRAFPAPFLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420

RESULT 15
US-08-748-591-9
; Sequence 9, Application US/08748591
; Patent No. 5759811
; GENERAL INFORMATION:
; APPLICANT: Epstein, Ervin
; APPLICANT: Hu, Zhilan
; APPLICANT: Bonifas, Jeanette
; TITLE OF INVENTION: Mutant Human Hedgehog Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish and Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,591
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06510/067001

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-748-591-9

Query Match      97.5%; Score 2407; DB 1; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.9e-232;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60
QY 61 RYEGKISRNSERFKELTPNPNPDIIPKDEBNTGADRLMTORCKDKLNALAI SVMNQPGV 120
DB 61 RYEGKISRNSERFKELTPNPNPDIIPKDEBNTGADRLMTORCKDKLNALAI SVMNQPGV 120
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180
DB 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
QY 241 FLDRDDGAKKVFYVETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
DB 241 FLDRDDGAKKVFYVETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAHVSVTLSEEAAGAYAPLTAQGTILINRVL 360
DB 301 PRALFASRVPRGQRYVVAERDGRLLPAHVSVTLSEEAAGAYAPLTAQGTILINRVL 360
QY 361 ASCVAVIEEHSWAHRAFPAPFLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420
DB 361 ASCVAVIEEHSWAHRAFPAPFLAHALLAALAPARTDRGSGGDRGGGGRVALTAPGA 420

RESULT 16
US-08-946-329A-20
; Sequence 20, Application US/08946329A
; Patent No. 6057091
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,329A
; FILING DATE: 07-OCT-1997

```

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/061,323  
 FILING DATE: 07-OCT-1996  
 APPLICATION NUMBER: 08/729,743  
 FILING DATE: 10-JUL-1996  
 APPLICATION NUMBER: 08/567,357  
 FILING DATE: 04-DEC-1995  
 APPLICATION NUMBER: 08/349,498  
 FILING DATE: 02-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: 07265/140001  
 TELEPHONE: 619/678-5070  
 TELEFAX: 619/678-5099  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 437 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-946-329A-20

Query Match 83.8%; Score 2068; DB 3; Length 437;  
 Best Local Similarity 87.4%; Pred. No. 1.6e-198;  
 Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

Qy	1	M L L A R C L L L V L S S L V C S G L A C G P G R G K R R H P K K L T P L A Y K Q F I P N V A E K T L G A S G	60
Db	2	L L L L A R C F L V I L A S S L L V C P L A C G P G R G K R R H P K K L T P L A Y K Q F I P N V A E K T L G A S G	61
Qy	61	R Y E G K I S R N S E R F K E L T P N Y N P D I I F K D E E N T G A D R L M T O R C K D K N A L A I S V M N Q W P G V	120
Db	62	R Y E G K I T R N S E R F K E L T P N Y N P D I I F K D E E N T G A D R L M T O R C K D K N A L A I S V M N Q W P G V	121
Qy	121	K L R V T E G W D E D G H S E S L H Y E G R A V D I T T S D R D R S K Y G M L A R L A V A G E F D W V Y E S K A H	180
Db	122	K L R V T E G W D E D G H S E S L H Y E G R A V D I T T S D R D R S K Y G M L A R L A V A G E F D W V Y E S K A H	181
Qy	181	I H C S V K A E N S V A A K S G G C F P G S A T V H L E Q G G T K L V K D L S P G D R V L A A D D Q G R L L Y S D F L T	240
Db	182	I H C S V K A E N S V A A K S G G C F P G S A T V H L E Q G G T K L V K D L S P G D R V L A A D D Q G R L L Y S D F L T	241
Qy	241	F L D R D G A K K V F Y I E T R P R E R L L T A A H L L F V A P H N D S A T G E P E A S G S G P P S G G A L G	300
Db	242	F L D R D G A K K V F Y I E T R P R E R L L T A A H L L F V A P H N D S A T G E P E A S G S G P P S G G A L G	286
Qy	301	P R A L F A S R V R P G Q R V Y V V A E R G D B R L L P A A V H S V T L S E E A G A Y A P L T A Q G T I L I N R V L	360
Db	287	P S A L F A S R V R P G Q R V Y V V A E R G D R R L L P A A V H S V T L R E E E A G A Y A P L T A H G T I L I N R V L	346
Qy	361	A S C Y A V I E E H S W A H R A F A P F R L A H A L A A L A P A R T D R G G D G G D R G G G G R V A L T A P G A	420
Db	347	A S C Y A V I E E H S W A H R A F A P F R L A H A L A A L A P A R T D R G G D G G D R G G G G R V A L T A P G A	395

RESULT 17  
 US-08-567-357A-20  
 Sequence 20, Application US/08567357A  
 Patent No. 6132728  
 GENERAL INFORMATION:  
 APPLICANT: Beachy, Philip A.  
 APPLICANT: Moon, Randall T.  
 APPLICANT: Porter, Jeffrey A.  
 TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400  
 CITY: La Jolla  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows 95  
 SOFTWARE: FastSeq for Windows Version 2.0b  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/567,357A  
 FILING DATE: 04-DEC-1995  
 APPLICATION DATA:  
 APPLICATION NUMBER: 08/349,498  
 FILING DATE: 02-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: 07265/080001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619/678-5070  
 TELEFAX: 619/678-5099  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 437 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-567-357A-20

Query Match 83.8%; Score 2068; DB 3; Length 437;  
 Best Local Similarity 87.4%; Pred. No. 1.6e-198;  
 Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

Qy	1	M L L A R C L L L V L S S L V C S G L A C G P G R G K R R H P K K L T P L A Y K Q F I P N V A E K T L G A S G	60
Db	2	L L L L A R C F L V I L A S S L L V C P L A C G P G R G K R R H P K K L T P L A Y K Q F I P N V A E K T L G A S G	61
Qy	61	R Y E G K I S R N S E R F K E L T P N Y N P D I I F K D E E N T G A D R L M T O R C K D K N A L A I S V M N Q W P G V	120
Db	62	R Y E G K I T R N S E R F K E L T P N Y N P D I I F K D E E N T G A D R L M T O R C K D K N A L A I S V M N Q W P G V	121
Qy	121	K L R V T E G W D E D G H S E S L H Y E G R A V D I T T S D R D R S K Y G M L A R L A V A G E F D W V Y E S K A H	180
Db	122	K L R V T E G W D E D G H S E S L H Y E G R A V D I T T S D R D R S K Y G M L A R L A V A G E F D W V Y E S K A H	181
Qy	181	I H C S V K A E N S V A A K S G G C F P G S A T V H L E Q G G T K L V K D L S P G D R V L A A D D Q G R L L Y S D F L T	240
Db	182	I H C S V K A E N S V A A K S G G C F P G S A T V H L E Q G G T K L V K D L S P G D R V L A A D D Q G R L L Y S D F L T	241
Qy	241	F L D R D G A K K V F Y I E T R P R E R L L T A A H L L F V A P H N D S A T G E P E A S G S G P P S G G A L G	300
Db	242	F L D R D G A K K V F Y I E T R P R E R L L T A A H L L F V A P H N D S A T G E P E A S G S G P P S G G A L G	286
Qy	301	P R A L F A S R V R P G Q R V Y V V A E R G D R R L L P A A V H S V T L S E E A G A Y A P L T A Q G T I L I N R V L	360
Db	287	P S A L F A S R V R P G Q R V Y V V A E R G D R R L L P A A V H S V T L R E E E A G A Y A P L T A H G T I L I N R V L	346
Qy	361	A S C Y A V I E E H S W A H R A F A P F R L A H A L A A L A P A R T D R G G D G G D R G G G G R V A L T A P G A	420
Db	347	A S C Y A V I E E H S W A H R A F A P F R L A H A L A A L A P A R T D R G G D G G D R G G G G R V A L T A P G A	395

RESULT 18  
 US-08-729-743A-20  
 Sequence 20, Application US/08729743A  
 Patent No. 6214794  
 GENERAL INFORMATION:

```

; APPLICANT: Beachy, Philip A.
; APPLICANT: Moon, Randall T.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA: US/08/729,743A
; FILING DATE: 07-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/099001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5039
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-729-743A-20

Query Match 83.8%; Score 2068; DB 3; Length 437;
Best Local Similarity 87.4%; Pred. No. 1.6e-198;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKKLTPLAYKQFIPNVAEKTGLGASG 60
DB 2 LLLARCFVLIVLASSLLVCPGLACGPGRGFKRRHPKKLTPLAYKQFIPNVAEKTGLGASG 61
QY 61 RYEGKISRNSRERFKELTNPYNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
DB 62 RYEGKITRNSRERFKELTNPYNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 121
QY 121 KLRVTEGWDEDDGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYVESKAH 180
DB 122 KLRVTEGWDEDDGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYVESKAH 181
QY 181 IHC5VKAENSVAAGSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 182 IHC5VKAENSVAAGSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDDGAKKVFYVETETPRERLLLTAAHLLFVAPHNDSATGEPASSGGSPGGALG 300
DB 242 FLDRDEGAKKVFYVETETPRERLLLTAAHLLFVAPHND-----SGTPG 286
QY 301 PRALPASVRVPGQVYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360
DB 287 PSALPASVRVPGQVYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAHGTILINRVL 346
QY 361 ASCYAVIEHSHWAHAFAPFRLAHALLAALAPARTDRCGDSGGDRGGGGGCVALTAPGA 420
DB 347 ASCYAVIEHSHWAHAFAPFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGTHWYSQLLYHICTWLLDSEALHPLGMVAKSS 462

; APPLICANT: Beachy, Philip A.
; APPLICANT: Moon, Randall T.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA: US/08/729,743A
; FILING DATE: 07-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/099001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5039
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-729-743A-20

Query Match 83.8%; Score 2068; DB 3; Length 437;
Best Local Similarity 87.4%; Pred. No. 1.6e-198;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKKLTPLAYKQFIPNVAEKTGLGASG 60
DB 2 LLLARCFVLIVLASSLLVCPGLACGPGRGFKRRHPKKLTPLAYKQFIPNVAEKTGLGASG 61
QY 61 RYEGKISRNSRERFKELTNPYNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
DB 62 RYEGKITRNSRERFKELTNPYNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 121
QY 121 KLRVTEGWDEDDGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYVESKAH 180
DB 122 KLRVTEGWDEDDGHSEESLHYEGRAVDITTSDDRSKYGMLARLAVAGFDWVYVESKAH 181
QY 181 IHC5VKAENSVAAGSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 182 IHC5VKAENSVAAGSGGCPGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDDGAKKVFYVETETPRERLLLTAAHLLFVAPHNDSATGEPASSGGSPGGALG 300
DB 242 FLDRDEGAKKVFYVETETPRERLLLTAAHLLFVAPHND-----SGTPG 286
QY 301 PRALPASVRVPGQVYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360
DB 287 PSALPASVRVPGQVYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAHGTILINRVL 346
QY 361 ASCYAVIEHSHWAHAFAPFRLAHALLAALAPARTDRCGDSGGDRGGGGGCVALTAPGA 420
DB 347 ASCYAVIEHSHWAHAFAPFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGTHWYSQLLYHICTWLLDSEALHPLGMVAKSS 462
```

QY 361 ASCYAVIEHSHWAHRAFAFFLAHALAALAPARTDRGGSGGGGRVVALTPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFFLAHALAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGTAGIHWSYQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGIHWSYQLLYQIGTWLLDSEALHPLGMVAVKSS 437

RESULT 20  
US-08-349-498-20  
; Sequence 20, Application US/08349498  
; Patent No. 6281332  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; APPLICANT: Moon, Randall T.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92037

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/349,498  
; FILING DATE: 02-DEC-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/043001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-349-498-20

Query Match 83.8%; Score 2068; DB 3; Length 437;  
Best Local Similarity 87.4%; Pred. No. 1.6e-198;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVSVLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCFVLIVASSLLVCFGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERFKELTPNPNPDIIFKDEBNTGADRLMTQRCCKDLNALAISVMNQWPGV 120  
Db 62 RYEGKITRNSERFKELTPNPNPDIIFKDEBNTGADRLMTQRCCKDLNALAISVMNQWPGV 121  
QY 121 KLRTVEGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 122 KLRTVEGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCYKVAENSVAAGSGCGCPGSGATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
Db 182 IHCYKVAENSVAAGSGCGCPGSGATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLVAPHNDSATGPEASGSGPPSGALG 300  
Db 242 FLDRDDGAKKVFYVETREPRERLLTAAHLLVAPHNDSATGPEASGSGPPSGALG 286

QY 301 PRALFASRVRCQRYVVAERDGRRLLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVRCQRYVVAERDGRRLLPAAVHVSVTLSSEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFAFFLAHALAALAPARTDRGGSGGGGRVVALTPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFFLAHALAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGTAGIHWSYQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGIHWSYQLLYQIGTWLLDSEALHPLGMVAVKSS 437

RESULT 21  
PCT-US95-15463-20  
; Sequence 20, Application PC/TUS9515463  
; GENERAL INFORMATION:  
; APPLICANT: The Johns Hopkins University School of Medicine  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/15463  
; FILING DATE: 01-DEC-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/080W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-15463-20

Query Match 83.8%; Score 2068; DB 5; Length 437;  
Best Local Similarity 87.4%; Pred. No. 1.6e-198;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVSVLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCFVLIVASSLLVCFGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERFKELTPNPNPDIIFKDEBNTGADRLMTQRCCKDLNALAISVMNQWPGV 120  
Db 62 RYEGKITRNSERFKELTPNPNPDIIFKDEBNTGADRLMTQRCCKDLNALAISVMNQWPGV 121  
QY 121 KLRTVEGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 122 KLRTVEGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCYKVAENSVAAGSGCGCPGSGATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
Db 182 IHCYKVAENSVAAGSGCGCPGSGATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLVAPHNDSATGPEASGSGPPSGALG 300

Db 242 FLDRDEGAKVFFVETLEPRERLILLTAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASRVPGQVYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPGQVYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAQGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTDRGGSGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWSOLLYQIGTWLDDSEALHPLGMVAKSS 462  
Db 396 TEARGAETAGIHWSOLLYQIGTWLDDSEALHPLGMVAKSS 437

RESULT 22  
PCT-US95-15923-20  
; Sequence 20, Application PC/TUS9515923  
; GENERAL INFORMATION:  
; APPLICANT: The Johns Hopkins University School of Medicine, et al.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/15923  
; FILING DATE: 04-DEC-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/043W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5039  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-15923-20

Query Match 83.8%; Score 2068; DB 5; Length 437;  
Best Local Similarity 87.4%; Pred. No. 1.6e-198;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

Qy 1 MILLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLFVILASLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
Qy 61 RYEGKISRNSERFKELTPNYPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQWPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQWPGV 121  
Qy 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYESKAH 180  
Db 122 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYESKAH 181  
Qy 181 IHCSYKAENSAKSGGCGPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSYKAENSAKSGGCGPGSATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241

Qy 241 FLDRDCGAKVFFVETLEPRERLILLTAHLLFVAPHNDSATCEPEASSGSGPSSGALG 300  
Db 242 FLDRDEGAKVFFVETLEPRERLILLTAHLLFVAPHND-----SGPTPG 286  
Qy 301 PRALFASRVPGQVYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPGQVYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAQGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTDRGGSGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPAPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWSOLLYQIGTWLDDSEALHPLGMVAKSS 462  
Db 396 TEARGAETAGIHWSOLLYQIGTWLDDSEALHPLGMVAKSS 437

RESULT 23  
US-08-176-427B-8  
; Sequence 8, Application US/08176427B  
; Patent No. 5789543  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/176,427B  
; FILING DATE: 30-DEC-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-176-427B-8

Query Match 83.6%; Score 2065; DB 1; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

Qy 1 MILLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLFVILASLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
Qy 61 RYEGKISRNSERFKELTPNYPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQWPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVNNQWPGV 121  
Qy 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWVYESKAH 180

Db 122 LRLVTEGWDGHHSEBSLHYEGRAVDITTSDDRSKYGMARLAVAEAGDWMVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERLLITAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASVRPGQVYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPGQVYVVAERGDRRLLPAAVHVSVTLEEEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFAFPRLAHALLAALAPARTDRGGSGGGDRGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFPRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWYSOLLYOIGTWLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGIHWYSOLLYHIGTWLLDSETMHPGLGMVAVKSS 437

RESULT 24  
US-08-356-060A-11  
; Sequence 11, Application US/08356060A  
; Patent No. 5844079  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,060A  
; FILING DATE: 14-DEC-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCES/DOCKET NUMBER: HMI-006CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-356-060A-11

Query Match 83.6%; Score 2065; DB 2; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3 3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLLVSSLLVCSGLACGFGKRRHPKLTPLAYKQFIPNVAEKTGLGASG 60

Db 2 LLLARCLVLLASSLLVCPGLACGFGKRRHPKLTPLAYKQFIPNVAEKTGLGASG 61  
QY 61 RYEGKISRNSRFRKELTNYNPDIIIFKDEENTGADRLMTQCKLNLALAI SVMNQPGV 120  
Db 62 RYEGKITRNSRFRKELTNYNPDIIIFKDEENTGADRLMTQCKLNLALAI SVMNQPGV 121  
QY 121 KLRTVEGWDGHHSEBSLHYEGRAVDITTSDDRSKYGMARLAVAEAGDWMVYVESKAH 180  
Db 122 LRLVTEGWDGHHSEBSLHYEGRAVDITTSDDRSKYGMARLAVAEAGDWMVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERLLITAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASVRPGQVYVVAERDGRRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPGQVYVVAERGDRRLLPAAVHVSVTLEEEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFAFPRLAHALLAALAPARTDRGGSGGGDRGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFPRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWYSOLLYOIGTWLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGIHWYSOLLYHIGTWLLDSETMHPGLGMVAVKSS 437

RESULT 25  
US-08-460-900C-11  
; Sequence 11, Application US/08460900C  
; Patent No. 6165747  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; APPLICANT: Bumcrot, David A.  
; APPLICANT: Marti-Gostiza, Elisa  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,900C  
; FILING DATE: 5-JUNE-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 4-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCES/DOCKET NUMBER: HMI-006.05

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-900C-11

Query Match 83.6%; Score 2065; DB 3; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60  
DB 2 LLLARCLFVLASSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQWPGV 120  
DB 62 RYEGKITRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQWPGV 121  
QY 121 KLAVTEGWDGDHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
DB 122 RLKRTGWDGDHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGRVLAADQGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGRVLAADQGRLLYSDFLT 241  
QY 241 FLORDGAKVYVYIETREPRERLLTAAHLLFVAPHNDSDATGEPEASSGSGPPSGGALG 300  
DB 242 FLORDGAKVYVYIETREPRERLLTAAHLLFVAPHNDSDATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSSEAGAYAPLTAQGTILINRVL 360  
DB 287 PSALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSSEAGAYAPLTAQGTILINRVL 346  
QY 361 ASCVAVIEHSWAHRAFPRLAHALLAALAPARTDGGSDGGGGRVALTAPGA 420  
DB 347 TEARGAEPAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
396 TEARGAEPAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 437

RESULT 26  
US-08-674-509B-11  
Sequence 11, Application US/08674509B  
Patent No. 6261786  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
APPLICANT: Mariago, Valeria  
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEHOG AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674,509B  
FILING DATE: 02-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/460,900  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-509B-11

Query Match 83.6%; Score 2065; DB 3; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60  
DB 2 LLLARCLFVLASSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQWPGV 120  
DB 62 RYEGKITRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQWPGV 121  
QY 121 KLAVTEGWDGDHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
DB 122 RLKRTGWDGDHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGRVLAADQGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGRVLAADQGRLLYSDFLT 241  
QY 241 FLORDGAKVYVYIETREPRERLLTAAHLLFVAPHNDSDATGEPEASSGSGPPSGGALG 300  
DB 242 FLORDGAKVYVYIETREPRERLLTAAHLLFVAPHNDSDATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSSEAGAYAPLTAQGTILINRVL 360  
DB 287 PSALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSSEAGAYAPLTAQGTILINRVL 346  
QY 361 ASCVAVIEHSWAHRAFPRLAHALLAALAPARTDGGSDGGGGRVALTAPGA 420  
DB 347 TEARGAEPAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
396 TEARGAEPAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSS 437

RESULT 27  
US-08-954-698-11  
Sequence 11, Application US/08954698  
Patent No. 6271363  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,698  
FILING DATE: 20-OCT-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-698-11

Query Match 83.6%; Score 2065; DB 3; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQIFPNVAEKTIGASG 60  
DB 2 LLLARCLVLTASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQIFPNVAEKTIGASG 61  
QY 61 RYEGKISRNSRFKELTPNPNDFIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
DB 62 RYEGKITRNSRFKELTPNPNDFIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 121  
QY 121 KLRTVEGWDEDDGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
DB 122 KLRTVEGWDEDDGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAKSGCGPFSATVHLEGGTKLVKOLSPGDRVLAADDGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAKSGCGPFSATVHLEGGTKLVKOLSPGDRVLAADDGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGFEASSGGSPSGGALG 300  
DB 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGTPPG 286  
QY 301 PRALFASVRPQORYVVAERDGRRLLPAAVHSTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 287 PSALFASVRPQORYVVAERDGRRLLPAAVHSTLSEEAAGAYAPLTAOGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFPFLAHALLAALAPARTDRGGDGGGGRGGRVALTAFGA 420  
DB 347 ASCYAVIEEHSWAHRAFPFLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395  
QY 421 ADAPAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMAVKSS 462  
DB 396 TEARGAEPATAGIHWYSQLLYHIGTWLLDSETHPLGMAVKSS 437

## RESULT 28

US-08-957-874-11  
Sequence 11, Application US/08957874  
Patent No. 6384192  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/957,874  
FILING DATE: 20-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.09  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-957-874-11

Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQIFPNVAEKTIGASG 60  
DB 2 LLLARCLVLTASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQIFPNVAEKTIGASG 61  
QY 61 RYEGKISRNSRFKELTPNPNDFIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
DB 62 RYEGKITRNSRFKELTPNPNDFIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 121  
QY 121 KLRTVEGWDEDDGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
DB 122 KLRTVEGWDEDDGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAKSGCGPFSATVHLEGGTKLVKOLSPGDRVLAADDGRLLYSDFLT 240

Db 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLRPGDRVLAADQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFVVIETREPRERLLLTAAHLLFVAPHNDSATGEPEASSGSGPSPSGALG 300  
Db 242 FLDRDEGAKKVFVVIETLEPRERLLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAPAFPRLAHALAALAPARTDGGDGGGGGGGGEVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAPAFPRLAHALAALAPARTD-----GGGGGIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWTLLDSEALHPLGMVAKSS 462  
Db 396 TEARGAEPTAGIHWSQLLYHIGTWTLLDSETMPLGMVAKSS 437  
RESULT 29  
US-09-325-256-20  
; Sequence 20, Application US/09325256  
; Patent No. 6444793  
; GENERAL INFORMATION:  
; APPLICANT: PEPINSKY, R. BLAKE  
; APPLICANT: BAKER, DARREN P.  
; APPLICANT: WEN, DINGYI  
; APPLICANT: WILLIAMS, KEVIN P.  
; APPLICANT: GARGER, ELLEN A.  
; APPLICANT: TAYLOR, FREDERICK R.  
; APPLICANT: GALDES, ALPHONSE  
; APPLICANT: PORTER, JEFFREY  
; TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND  
; FILE REFERENCE: BIV-067.01  
; CURRENT APPLICATION NUMBER: US/09/325,256  
; CURRENT FILING DATE: 1999-06-03  
; PRIOR APPLICATION NUMBER: 60/099,800  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/089,935  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/067,423  
; PRIOR FILING DATE: 1997-12-03  
; PRIOR APPLICATION NUMBER: PCT/US98/25676  
; PRIOR FILING DATE: 1998-12-03  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Murine sp.  
US-09-325-256-20  
Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLILARCLLVLVSSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTLCASG 60  
Db 2 LLLILARCLVILASLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTLCASG 61  
QY 61 RYEGKISRNSRPFKELTPNYPDIIFKDEBNTGADRLMTORCKDKNALAISVYNQWPGV 120  
Db 62 RYEGKITRNSRPFKELTPNYPDIIFKDEBNTGADRLMTORCKDKNALAISVYNQWPGV 121  
QY 121 KLRVTEGWDEGHSESLHVEGRAVDITTSDDRSKYGLARLAVBAGFDWVYIESKAH 180  
Db 122 KLRVTEGWDEGHSESLHVEGRAVDITTSDDRSKYGLARLAVBAGFDWVYIESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLRPGDRVLAADQGRLLYSDFLT 240

Db 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLRPGDRVLAADQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFVVIETREPRERLLLTAAHLLFVAPHNDSATGEPEASSGSGPSPSGALG 300  
Db 242 FLDRDEGAKKVFVVIETLEPRERLLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAPAFPRLAHALAALAPARTDGGDGGGGGGGGEVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAPAFPRLAHALAALAPARTD-----GGGGGIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWTLLDSEALHPLGMVAKSS 462  
Db 396 TEARGAEPTAGIHWSQLLYHIGTWTLLDSETMPLGMVAKSS 437  
RESULT 30  
US-09-639-695-11  
; Sequence 11, Application US/09639695  
; Patent No. 6576237  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; McMahon, Andrew P.  
; Tabin, Clifford J.  
; Bumcrot, David A.  
; Marti-Gorostiza, Elisa  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/639,695  
; FILING DATE: 16-Aug-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 4-MAY-1995  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMV-006.05  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-639-695-11  
Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MILLARCLLLVVSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 2 LLLARCLVTLASLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 61  
QY 61 RYEGKISRNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
Db 62 RYEGKITNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVMNQWPGV 121  
QY 121 KLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Db 122 KLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 181  
QY 181 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLRFGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALPASVRPGQVYVVAERDGRRLLPAAVHVSVTLSSEAAAGAYAPLTAOGTILINRVL 360  
Db 287 PSALPASVRPGQVYVVAERGDRRLLPAAVHVSVTLEEEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGATAGHWSQLLYOIGTWTLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGHWSQLLYHIGTWTLLDSETHWHPLOMAVKSS 437  
RESULT 31  
US-09-448-188-11  
; Sequence 11, Application US/09448188  
; Patent No. 6607913  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; McMahon, Andrew P.  
; Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/448.188  
; FILING DATE: 23-NOV-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/462,386  
; FILING DATE: 05-JUN-1995  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 04-MAY-1995  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMV-006.12

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-448-188-11  
Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MILLARCLLLVVSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 2 LLLARCLVTLASLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 61  
QY 61 RYEGKISRNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
Db 62 RYEGKITNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAIISVMNQWPGV 121  
QY 121 KLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Db 122 KLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 181  
QY 181 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLRFGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALPASVRPGQVYVVAERDGRRLLPAAVHVSVTLSSEAAAGAYAPLTAOGTILINRVL 360  
Db 287 PSALPASVRPGQVYVVAERGDRRLLPAAVHVSVTLEEEAGAYAPLTAHGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGATAGHWSQLLYOIGTWTLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAETAGHWSQLLYHIGTWTLLDSETHWHPLOMAVKSS 437  
RESULT 32  
US-08-954-128-11  
; Sequence 11, Application US/08954128  
; Patent No. 6610656  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; McMahon, Andrew P.  
; Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/954.128  
; FILING DATE: 23-NOV-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/462,386  
; FILING DATE: 05-JUN-1995  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 04-MAY-1995  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMV-006.12

APPLICATION NUMBER: US/08/954,128  
FILING DATE: 20-OCT-1997  
PRIORITY APPLICATION DATA: US 08/462,386  
PRIORITY APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
PRIORITY APPLICATION DATA: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-128-11

Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198; Indels 26; Gaps 3;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTGASG 60  
DB 2 LLLARCLFVLASSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTGASG 61  
QY 61 RYEGKISNSERPKELTPNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNWQPGV 120  
DB 62 RYEGKITRNSERPKELTPNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNWQPGV 121  
QY 121 KLRTVEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
DB 122 LRLVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGGTVKLDLSPGDRVLAADDOGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAGSGCGFPGSATVHLEOGGTVKLDLSPGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDGAKKVYVYIETREPRERLLTAHLLFVAPHNDSATGEPEASGSGPPSGGALG 300  
DB 242 FLDRDGAKKVYVYIETREPRERLLTAHLLFVAPHNDSATGEPEASGSGPPSGGALG 300  
QY 301 PRALFASVRPQGVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 287 PSALFASVRPQGVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRVL 346  
QY 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDGGSGGDRGGRGGRVALTAPGA 420  
DB 347 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDGGSGGDRGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
DB 396 TEARGABPTAGIHWSQLLYHIGTWLLDSETHMPLGMVAVKSS 437

RESULT 33  
US-09-704-917-13  
Sequence 13, Application US/09704917  
Patent No. 6616926  
GENERAL INFORMATION:  
APPLICANT: Biogen, Inc.  
APPLICANT: Burkly, Linda

APPLICANT: Wang, Li Chun  
TITLE OF INVENTION: METHODS OF MODULATING LIPID METABOLISM AND STORAGE  
FILE REFERENCE: A069PCT  
CURRENT APPLICATION NUMBER: US/09/704,917  
CURRENT FILING DATE: 2000-11-02  
PRIOR APPLICATION NUMBER: 60/122,640  
PRIOR FILING DATE: 1999-03-03  
PRIOR APPLICATION NUMBER: 60/124,446  
PRIOR FILING DATE: 1999-03-15  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 13  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Murine sp.  
US-09-704-917-13  
Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198; Indels 26; Gaps 3;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTGASG 60  
DB 2 LLLARCLFVLASSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTGASG 61  
QY 61 RYEGKISNSERPKELTPNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNWQPGV 120  
DB 62 RYEGKITRNSERPKELTPNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNWQPGV 121  
QY 121 KLRTVEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
DB 122 LRLVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEOGGTVKLDLSPGDRVLAADDOGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAGSGCGFPGSATVHLEOGGTVKLDLSPGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDGAKKVYVYIETREPRERLLTAHLLFVAPHNDSATGEPEASGSGPPSGGALG 300  
DB 242 FLDRDGAKKVYVYIETREPRERLLTAHLLFVAPHNDSATGEPEASGSGPPSGGALG 300  
QY 301 PRALFASVRPQGVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 287 PSALFASVRPQGVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRVL 346  
QY 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDGGSGGDRGGRGGRVALTAPGA 420  
DB 347 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDGGSGGDRGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
DB 396 TEARGABPTAGIHWSQLLYHIGTWLLDSETHMPLGMVAVKSS 437  
RESULT 34  
US-08-954-740-11  
Sequence 11, Application US/08954740  
Patent No. 6630148  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:



STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICANT: de Sauvage, Frederic  
TITLE OF INVENTION: Patched-2  
FILE REFERENCE: P140SR1  
CURRENT APPLICATION NUMBER: US/09/293,505  
CURRENT FILING DATE: 1999-04-15  
EARLIER APPLICATION NUMBER: US 60/081,884  
EARLIER FILING DATE: 1998-04-15  
NUMBER OF SEQ ID NOS: 32  
SEQ ID NO 14  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-293-505-14  
Query Match 83.5%; Score 2062; DB 4; Length 437;  
Best Local Similarity 87.0%; Pred. No. 6.5e-198;  
Matches 402; Conservative 10; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCFVLVILASSLLVCPGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 121  
QY 121 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYIESKAH 180  
Db 122 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYIESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFFGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGCGFFGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAVHSVTLSSEAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQVYVVAERDGRLLPAVHSVTLSSEAGAYAPLTAQGTILINRVL 346  
QY 361 ASCVAVIEEHSWAHRAFPRLAHALLAALAPARTDRGSGGDRGGGRVALTPGA 420  
Db 347 ASCVAVIEEHSWAHRAFPRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWYSOLLVIGTWLLDSBALHPLGNVAVKSS 462  
Db 396 TEARGAETAGIHWYSOLLVIGTWLLDSETMPLGNVAVKAS 437  
RESULT 37  
US-08-757-230A-2  
Sequence 2, Application US/08757230A  
Patent No. 6235885  
GENERAL INFORMATION:  
APPLICANT: Thomas M. Jessell, et al.  
TITLE OF INVENTION: RAT HEDGEHOG PROTEIN-1 (VHH-1)  
FILE REFERENCE: 0575/45375-1/JPM/SHS/MVM  
CURRENT APPLICATION NUMBER: US/08/757,230A  
CURRENT FILING DATE: 1996-11-27  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 437  
TYPE: PRT  
ORGANISM: RAT  
US-08-757-230A-2

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICANT: de Sauvage, Frederic  
TITLE OF INVENTION: Patched-2  
FILE REFERENCE: P140SR1  
CURRENT APPLICATION NUMBER: US/09/293,505  
CURRENT FILING DATE: 1999-04-15  
EARLIER APPLICATION NUMBER: US 60/081,884  
EARLIER FILING DATE: 1998-04-15  
NUMBER OF SEQ ID NOS: 32  
SEQ ID NO 14  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-293-505-14  
Query Match 83.6%; Score 2065; DB 4; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.3e-198;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCFVLVILASSLLVCPGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 121  
QY 121 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYIESKAH 180  
Db 122 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYIESKAH 181  
QY 181 IHCSVKAENSVAAGSGCGFFGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGCGFFGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300  
Db 242 FLDRDEGAKKVFYVETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAVHSVTLSSEAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQVYVVAERDGRLLPAVHSVTLSSEAGAYAPLTAQGTILINRVL 346  
QY 361 ASCVAVIEEHSWAHRAFPRLAHALLAALAPARTDRGSGGDRGGGRVALTPGA 420  
Db 347 ASCVAVIEEHSWAHRAFPRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWYSOLLVIGTWLLDSBALHPLGNVAVKSS 462  
Db 396 TEARGAETAGIHWYSOLLVIGTWLLDSETMPLGNVAVKAS 437  
RESULT 37  
US-08-757-230A-2  
Sequence 2, Application US/08757230A  
Patent No. 6235885  
GENERAL INFORMATION:  
APPLICANT: Thomas M. Jessell, et al.  
TITLE OF INVENTION: RAT HEDGEHOG PROTEIN-1 (VHH-1)  
FILE REFERENCE: 0575/45375-1/JPM/SHS/MVM  
CURRENT APPLICATION NUMBER: US/08/757,230A  
CURRENT FILING DATE: 1996-11-27  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 437  
TYPE: PRT  
ORGANISM: RAT  
US-08-757-230A-2

Query Match	83.4%; Score 2060; DB 3; Length 437;
Best Local Similarity	87.5%; Pred. No. le-197;
Matches 405; Conservative 8; Mismatches 22; Indels 28; Gaps 4;	
QY	1 MLLARCLLLVLVSSLLVCSGLACPGRGFGKGRHPKKLTPLAYKQFIPNVAEKTLCASG 60
DB	: : LLLARCFVALASSLLVCGLACPGRGFGKQHPPKLTPLAYKQFIPNVAEKTLCASG 61 :
QY	61 RYEGKISRNSRFKELTPNTNPDIIFKDENTGADRLMTQRCKDKLNALAI SVMNQPGV 120
DB	: : RYEGKITRNSRFKELTPNTNPDIIFKDENTGADRLMTQRCKDKLNALAI SVMNQPGV 121 :
QY	121 KLRVTGEWDEGDGHSESLHYEGRAVDITTTSDRDRSKYGMLARLAVEAGFDVVYYESKAH 180
DB	: : KLRVTGEWDEGDGHSESLHYEGRAVDITTTSDRDRSKYGMLAE LAVEAGFDVVYYESKAR 181 :
QY	181 IHC SKAENSVAAKSGGGCPGOSATVHL EOGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240
DB	: : IHC SKAENSVAAKSGDGC PPGSATVHLEOGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 241 :
QY	241 FLDRDDGAKVPFYVIETREPRERLLTAAHLLVFAPHND SATGEPEASSGSGPSPGGGALG 300
DB	: : FLDRREGAKVPFYVIETREPRERLLL TAAHLL VFAPHND-----SGPTPG 286 :
QY	301 PRALFASRVRCQRYVVVVAEDGBDRRLLPAAVHSVTLSEBAAGAYAPLTAQTILLINRVL 360
DB	: : PSPLFASRVRCQRYVVVVAERGGBRRLLPAAVHSVTLREBAAGAYAPLTAQTILLINRVL 346 :
QY	361 ASCYAVIEHSHWAHRAPAFPFLAHALLAALAPARTDGGSGDGGGGGGGVALTAP-G 419
DB	: : ASCYAVIEHSHWAHRAPAFPFLAHALLAALAPARTD-----GGGGG--SIFAPQS 394 :
QY	420 AADAPGAGATAGIHYSOLLYOIGTWLLDSEALUHPIGNVKSS 462
DB	: : VAEARGAGPAGIHHYSOLLYHIGTWLLDSETLHPGMVKSS 437 :

```

RESULT 39
US-08-757-230A-9
; Sequence 9, Application US/08757230A
; Patent No. 6235885
; GENERAL INFORMATION:
; APPLICANT: Thomas M. Jessell, et al.
; TITLE OF INVENTION: RAT HEDGEHOG PROTEIN-1 (VHH-1)
; FILE REFERENCE: 0575/45375-1/JPW/SHS/WM
; CURRENT APPLICATION NUMBER: US/06/757,230A
; CURRENT FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Rat
US-08-757-230A-9

Query Match      83.4%; Score 2060; DB 3; Length 437;
Best Local Similarity 87.5%; Pred. No. le-197;
Matches 405; Conservative 8; Mismatches 22; Indels 28; Gaps 4

Qy      1  MLLARCLLVLVSSILVCSGLACPGRGFKBHPKKLPLAYKQIPNVAEKTLCAGS 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2  LLLARCFVLVALASSLLVCFGLACPGRGFKBHPKKLPLAYKQIPNVAEKTLCAGS 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      61  RYEGKISRNSERAFKELTPNPNDDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      62  RYEGKITRNSERAFKELTPNPNDDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPGV 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      121  KLRVTEGWDGHHSEESHVYSGRAVDITTTSDRDSKYGMRLARLAVEAGPDWYVYESKAH 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      122  KLRVTEGWDGHHSEESHVYSGRAVDITTTSDRDSKYGMRLARLAVEAGPDWYVYESKAH 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      181  IHC5VKAENSAVAAKSGCGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDGCRLLYSDFLT 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Query Match	83.4%; Score 2060; DB 4; Length 437;
Best Local Similarity	87.5%; Pred. No. 1e-197;
Matches	405; Conservative 8; Mismatches 22; Indels 28; Gaps 4
Qy	1 MLLARCLLVLSGLVCSGLACGPGFGKRRHPKKLTPLAYKQFIPNVAEKTGLGSG 60
Db	2 LLLARCFVLALASLLVCPGLACGPGFGKRRHPKKLTPLAYKQFIPNVAEKTGLGSG 61
Qy	61 RYEGKISRNSRFRKELTPNYPDIIFKDEBNTGADRLMTQCKDKLNALAI SVMNQPGV 120
Db	62 RYEGKITRNSRFRKELTPNYPDIIFKDEBNTGADRLMTQCKDKLNALAI SVMNQPGV 121
Qy	121 KLRVTEGDEGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVEAGFQWVYVESKAH 180
Db	122 KLRVTEGDEGHHSESLHYEGRAVDITTSDRRSKYGMRLARLAVEAGFQWVYVESKAR 181
Qy	181 IHCSVKAENSVAAXSGGCFPGSATVHLHQGGTKLVKDLSPGRVLAAADQGRLLYSDFLT 240
Db	182 IHCSVKAENSVAAXSDGCFPGSATVHLHQGGTKLVKDLSPGRVLAAADQGRLLYSDFLT 241
Qy	241 FLDRDDGAKVYVIETREPRERILLTAAHLLFVAPHNDSATGPEASSGGSPSGGALG 300
Db	242 FLDRDEGAKVYVIETREPRERILLTAAHLLFVAPHND-----SGPTPG 286
Qy	301 PRALPASVRPQGVVYVAERDGRRLPPAAVHSVTLSEEAAGAVAPLTAGOTILINRVL 360
Db	287 PSPLPASVRPQGVVYVAERGGDRLLPAAVHSVTLREEAAGAVAPLTADGOTILINRVL 346
Qy	361 ASCYAVITEHSHWARAFAPFLHALALAAALAPARTDRGDSGGDGGGGGRGVALTAP-G-419
Db	347 ASCYAVITEHSHWARAFAPFLHALALAAALAPARTD-----GGGGG--SIPASQS 394

QY 420 AADAPGAGATAGIHWYSQLLYQIGTWLIDSEALHPLGMVAKSS 462  
Db 395 VAEARGAGPPAGIHWYSQLLYHIGTWLIDSETLHPLGMVAKSS 437

RESULT 41  
PCT-US95-02315-2  
; Sequence 2, Application PC/TUS9502315  
; GENERAL INFORMATION:  
; APPLICANT: Jessell, Thomas M.  
; APPLICANT: Dodd, Jane  
; APPLICANT: Roelink, Henk  
; APPLICANT: Edlund, Thomas  
; TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOG OF  
; TITLE OF INVENTION: HDGSHOG, VHH-1, EXPRESSED BY THE NOTOCHORD, AND USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02315  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: John P. White  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 45375-A-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-02315-2

Query Match 83.4%; Score 2060; DB 5; Length 437;  
Best Local Similarity 87.5%; Pred. No. 1e-197;  
Matches 405; Conservative 8; Mismatches 22; Indels 28; Gaps 4;  
QY 1 MLLARCLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 60  
Db 2 LLLARCLFVALASSLLVCPGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 61  
QY 61 RYEGKISRNSRFKELTNPYNDIIFKDEENTGADRLMTQCKDKNALAISVNNQWPGV 120  
Db 62 RYEGKITNSRFKELTNPYNDIIFKDEENTGADRLMTQCKDKNALAISVNNQWPGV 121  
QY 121 KLRVTEGDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 122 KLRVTEGDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAR 181  
QY 181 IHCSVKAENSVAAGSGGCPGSAIVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGGCPGSAIVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKVYFYIETREPRERLLLTAAHLLFVAPHNDSATGPEPEASSGGPPSGGALG 300  
Db 242 FLDRDEGAKVYFYIETREPRERLLLTAAHLLFVAPHND-----SGPTPG 286

QY 301 PRALFASRVPRGORVYVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSPLFASRVPRGORVYVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFAPFRLAHALLAALAPARTDGGSDGGGGRVALTAP-G 419  
Db 347 ASCYAVIEEHSWAHRAFAPFRLAHALLAALAPARTD-----GGGGG--SIPAPQS 394  
QY 420 AADAPGAGATAGIHWYSQLLYQIGTWLIDSEALHPLGMVAKSS 462  
Db 395 VAEARGAGPPAGIHWYSQLLYHIGTWLIDSETLHPLGMVAKSS 437

RESULT 42  
US-08-176-427B-2  
; Sequence 2, Application US/08176427B  
; Patent No. 5789543  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/176,427B  
; FILING DATE: 30-DEC-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 425 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-176-427B-2

Query Match 73.2%; Score 1807; DB 1; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;  
QY 1 MLLARCLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 60  
Db 4 MLLTRILLVGFICALLVSSGLTCGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 63  
QY 61 RYEGKISRNSRFKELTNPYNDIIFKDEENTGADRLMTQCKDKNALAISVNNQWPGV 120  
Db 64 RYEGKITNSRFKELTNPYNDIIFKDEENTGADRLMTQCKDKNALAISVNNQWPGV 123  
QY 121 KLRVTEGDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 124 KLRVTEGDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 183  
QY 181 IHCSVKAENSVAAGSGGCPGSAIVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

Db 184 IHCSVKAENSVAAGSGCGFPGSATVHLEHGGTKLVKDLSPGDRVLAADADGRLLYSDFLT 243  
Qy 241 FLDRDDGAKKVFYVYIETREPRERLLITAAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA 298  
Db 244 FLDRMDSRKLFFYIETROPRLRLTAAHLLFVAPQHNOSEATG-----STSG----- 292  
Qy 299 LGPRALFASRVPRQORVYVVAERDGRRLPAAVHVSVTLSZEAAGAYAPLTAQGTILINR 358  
Db 293 ---QALFASNVKFGQRYVVLGE--GGQQLLPASVHVSLSREASGAYAPLTAQGTILINR 347  
Qy 359 VLASCVAVIEHSHWAHAFAPFRLAHALLAALAPARTDRGDSGGDRGGGGRGVALTAP 418  
Db 348 VLASCVAVIEHSHWAHAFAPFRLAQGLLAL-----CP 381  
Qy 419 GAADAPGAGATAGHWTSQLLYQIGTWLDDSEALHPLGMVAKSS 462  
Db 382 DGAIPTAATTTTGIHWSRLLYRIGSVWLDGDLHPLGMVAPAS 425

RESULT 43  
US-08-356-060A-8  
; Sequence 8, Application US/08356060A  
; Patent No. 5844079  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,060A  
; FILING DATE: 14-DEC-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-006CP  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 425 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-356-060A-8

Query Match 73.2%; Score 1807; DB 2; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;  
Qy 1 MLLARCLLVLVSSLLVCSGLACGFGRGFRHPKLTPLAYKQFIPNVAEKTGLGSG 60  
Db 4 MLLTRILVGVFCALLVSSGLTCGFRGIGKRHPKLTPLAYKQFIPNVAEKTGLGSG 63  
Qy 61 RYEGKISRNSRFRKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAI-SVMNQPGV 120

Db 64 RYEGKISRNSRFRKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAI-SVMNQPGV 123  
Qy 121 KLRVTEGDEGHSSELSHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 124 KLRVTEGDEGHSSELSHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 183  
Qy 181 IHCSVKAENSVAAGSGCGFPGSATVHLEHGGTKLVKDLSPGDRVLAADADGRLLYSDFLT 240  
Db 184 IHCSVKAENSVAAGSGCGFPGSATVHLEHGGTKLVKDLSPGDRVLAADADGRLLYSDFLT 243  
Qy 241 FLDRDDGAKKVFYVYIETREPRERLLITAAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA 298  
Db 244 FLDRMDSRKLFFYIETROPRLRLTAAHLLFVAPQHNOSEATG-----STSG----- 292  
Qy 299 LGPRALFASRVPRQORVYVVAERDGRRLPAAVHVSVTLSZEAAGAYAPLTAQGTILINR 358  
Db 293 ---QALFASNVKFGQRYVVLGE--GGQQLLPASVHVSLSREASGAYAPLTAQGTILINR 347  
Qy 359 VLASCVAVIEHSHWAHAFAPFRLAHALLAALAPARTDRGDSGGDRGGGGRGVALTAP 418  
Db 348 VLASCVAVIEHSHWAHAFAPFRLAQGLLAL-----CP 381  
Qy 419 GAADAPGAGATAGHWTSQLLYQIGTWLDDSEALHPLGMVAKSS 462  
Db 382 DGAIPTAATTTTGIHWSRLLYRIGSVWLDGDLHPLGMVAPAS 425

RESULT 44  
US-08-460-900C-8  
; Sequence 8, Application US/08460900C  
; Patent No. 6165747  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tabin, Clifford J.  
; APPLICANT: Bumcirot, David A.  
; APPLICANT: Marti-Gorostiza, Elisa  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLSEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,900C  
; FILING DATE: 5-JUNE-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 4-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/356,060  
; FILING DATE: 14-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMV-006.05  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-900C-8

Query Match 73.2%; Score 1807; DB 3; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 1 MLLARCLLVSVLLVSGLAGCGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
DB 4 MLLTRILLVGFICALLVSSGLTCGPGRGIGKRRHPKLTPLAYKQFIPNVAEKTIGASG 63

QY 61 RYEGKISRNSERFKELTPNYPDIIPKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 120  
DB 64 RYEGKITRNSERFKELTPNYPDIIPKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 123

QY 121 KLRVTEGDEDDHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180  
DB 124 KLRVTEGDEDDHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 183

QY 181 IHCSVKAENSVAKSGCGPFGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 184 IHCSVKAENSVAKSGCGPFGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 243

QY 241 FLDRDDGAKVFFVIETREPRERLLTAAHLFVAP-HNDS-ATGEPEASSGGPPSGGA 298  
DB 244 FLDRMSSRKLFFVIETROPRLTAAHLFVAPQHNQSEATG-----STSG----- 292

QY 299 LGPRALPASVRPQGVVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAAQTILINR 358  
DB 293 ---QALFASNVPQGVVVLGE---GGQQLLPASVHSVLSREESAGAYAPLTAAQTILINR 347

QY 359 VLASCYAVIEHSHWAHRAFPFLAHLAALAPARTDRGSDGGGGRGGRVALTAP 418  
DB 348 VLASCYAVIEHSHWAHRAFPFLAHLAALAPARTDRGSDGGGGRGGRVALTAP 418

QY 419 GAADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSS 462  
DB 382 DGAIPATAATTTGTHWYSRLLYRIGSWVLDGDLHPLGMVAPAS 425

RESULT 45  
US-08-674-509B-8  
Sequence 8, Application US/08674509B  
Patent No. 6261786  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
APPLICANT: Marigo, Valeria  
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEHOG AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/674,509B  
FILING DATE: 02-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/460,900  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.06  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-509B-8

Query Match 73.2%; Score 1807; DB 3; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 1 MLLARCLLVSVLLVSGLAGCGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
DB 4 MLLTRILLVGFICALLVSSGLTCGPGRGIGKRRHPKLTPLAYKQFIPNVAEKTIGASG 63

QY 61 RYEGKISRNSERFKELTPNYPDIIPKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 120  
DB 64 RYEGKITRNSERFKELTPNYPDIIPKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 123

QY 121 KLRVTEGDEDDHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180  
DB 124 KLRVTEGDEDDHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 183

QY 181 IHCSVKAENSVAKSGCGPFGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 184 IHCSVKAENSVAKSGCGPFGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 243

QY 241 FLDRDDGAKVFFVIETREPRERLLTAAHLFVAP-HNDS-ATGEPEASSGGPPSGGA 298  
DB 244 FLDRMSSRKLFFVIETROPRLTAAHLFVAPQHNQSEATG-----STSG----- 292

QY 299 LGPRALPASVRPQGVVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAAQTILINR 358  
DB 293 ---QALFASNVPQGVVVLGE---GGQQLLPASVHSVLSREESAGAYAPLTAAQTILINR 347

QY 359 VLASCYAVIEHSHWAHRAFPFLAHLAALAPARTDRGSDGGGGRGGRVALTAP 418  
DB 348 VLASCYAVIEHSHWAHRAFPFLAHLAALAPARTDRGSDGGGGRGGRVALTAP 418

QY 419 GAADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSS 462  
DB 382 DGAIPATAATTTGTHWYSRLLYRIGSWVLDGDLHPLGMVAPAS 425

RESULT 46  
US-08-954-698-8  
Sequence 8, Application US/08954698  
Patent No. 6271363  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
APPLICANT: Marigo, Valeria  
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEHOG AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,698  
FILING DATE: 02-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

Sequence 8, Application US/08957874  
Patent No. 6384192  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/957,874  
FILING DATE: 20-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.09  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-957-874-8

Query Match 73.2%; Score 1807; DB 4; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

Qy	1	MLLARCLLLVLSVSLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG	60
Db	4	MLLTTRILLVGFICALLVSSGLTCGPGRGIGKRRHPKLTPLAYKQFIPNVAEKTIGASG	63
Qy	61	RYEGKISRNSRPFKELTNPYNDIIFKDEBENTGADRLMTQCKDKLNALAISVMNQPGV	120
Db	64	RYEGKITRNSRPFKELTNPYNDIIFKDEBENTGADRLMTQCKDKLNALAISVMNQPGV	123
Qy	121	KLRVTGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH	180
Db	124	KLRVTGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH	183
Qy	181	IHCSVKAENSVAAKSGCGPFGSATVHLBOGTKLVKDLSPGDRVLAADQGRLLYSDFLT	240
Db	184	IHCSVKAENSVAAKSGCGPFGSATVHLBOGTKLVKDLSPGDRVLAADQGRLLYSDFLT	243
Qy	241	FLDRDDGAKKVFYIETREPRERLLLTAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA	298

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,698  
FILING DATE: 20-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.10  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-698-8

Query Match 73.2%; Score 1807; DB 3; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

Qy	1	MLLARCLLLVLSVSLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG	60
Db	4	MLLTTRILLVGFICALLVSSGLTCGPGRGIGKRRHPKLTPLAYKQFIPNVAEKTIGASG	63
Qy	61	RYEGKISRNSRPFKELTNPYNDIIFKDEBENTGADRLMTQCKDKLNALAISVMNQPGV	120
Db	64	RYEGKITRNSRPFKELTNPYNDIIFKDEBENTGADRLMTQCKDKLNALAISVMNQPGV	123
Qy	121	KLRVTGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH	180
Db	124	KLRVTGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYSKAH	183
Qy	181	IHCSVKAENSVAAKSGCGPFGSATVHLBOGTKLVKDLSPGDRVLAADQGRLLYSDFLT	240
Db	184	IHCSVKAENSVAAKSGCGPFGSATVHLBOGTKLVKDLSPGDRVLAADQGRLLYSDFLT	243
Qy	241	FLDRDDGAKKVFYIETREPRERLLLTAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA	298
Db	244	FLDRDSSGKLFYIETREPRERLLLTAHLLFVAPQHNQSEATG-----STSG-----	292
Qy	299	LQPRALFASRVPFGORVVVAERDGRLLPAVHSLTSEAGAYAPLTAQGTILINR	358
Db	293	---QALFASNVKPGORVVYVLE--GGQQLLPASVHSLTSEAGAYAPLTAQGTILINR	347
Qy	359	VLASCYAVITEERSWAHRAPFRLAHLAALAPARTDGGDGGGGRGGRVALTAP	418
Db	348	VLASCYAVITEERSWAHRAPFRLAHLAALAPARTDGGDGGGGRGGRVALTAP	381
Qy	419	GAADPAGCATAGIHWYSLLYIGTWTLLDSEALHPLGNVYKSS	462
Db	382	DGAIPATAATTTGIHWYSLLYIGTWTLLDSEALHPLGNVYKSS	425

RESULT 47  
US-08-957-874-8

Db 244 FLDRMSSRKLFFVETROPARLLTAAHLLFVAPQHNSQSEATG-----STSG----- 292

Qy 299 LGPRALFASRVPRGQVYVVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILNR 358

Db 293 ---QALFASNVKPGQVYVLGE--GGQQLLPASVHSVLSREASGAYAPLTAQGTILNR 347

Qy 359 VLASCYAVIEHSHWAHRAFAFAPFLAHALLAALAPARTDRGDSGGGDRGGGGRVALTAP 418

Db 348 VLASCYAVIEHSHWAHRAFAFAPFLAQLLAAL-----CP 381

Qy 419 GAADAPCAGATAGIHWSYQLLYOIGTWLSDSEALHPLGMVAVKSS 462

Db 382 DGAIPTAATTTGTHWYSRLLYRIGSWLVDGDLHPLGMVAVPAS 425

RESULT 48

US-09-325-256-17

; Sequence 17, Application US/09325256

; Patent No. 6444793

; GENERAL INFORMATION:

; APPLICANT: PEPINSKY, R. BLAKE

; APPLICANT: BAKER, DARREN P.

; APPLICANT: WEN, DINGYI

; APPLICANT: WILLIAMS, KEVIN P.

; APPLICANT: GARGER, ELLEN A.

; APPLICANT: TAYLOR, FREDERICK R.

; APPLICANT: GALDES, ALPHONSE

; APPLICANT: PORTER, JEFFREY

; TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND

; FILE REFERENCE: BIV-067.01

; CURRENT APPLICATION NUMBER: US/09/325,256

; PRIOR FILING DATE: 1999-06-03

; PRIOR APPLICATION NUMBER: 60/099,800

; PRIOR FILING DATE: 1998-09-10

; PRIOR APPLICATION NUMBER: 60/078,935

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/089,685

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/067,423

; PRIOR FILING DATE: 1997-12-03

; PRIOR APPLICATION NUMBER: PCT/US98/25676

; PRIOR FILING DATE: 1998-12-03

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 17

; LENGTH: 425

; TYPE: PRT

; ORGANISM: Gallus sp.

US-09-325-256-17

Query Match 73.2%; Score 1807; DB 4; Length 425;

Best Local Similarity 77.8%; Pred. No. 2.2e-172;

Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

Qy 1 MLLARCLLVLSLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 60

Db 4 MLLTRILLVGFICALLVSSGLTCGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 63

Qy 61 RYEGKISRNSRRKELTPNPNPDIIFKDEENTGADRLMTORCKKLNALAISVWQWPGV 120

Db 64 RYEGKITRNSRRKELTPNPNPDIIFKDEENTGADRLMTORCKKLNALAISVWQWPGV 123

Qy 121 KLRVTEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGFDWYVESKAH 180

Db 124 KLRVTEGWDEGHHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGFDWYVESKAH 183

Qy 181 IHCSVKAENSVAAGSGGCPGASATVHLEOGTKLVKDLSPDRVLAADDGRLLYSDFLT 240

Db 184 IHCSVKAENSVAAGSGGCPGASATVHLEHGGTKLVKDLSPDRVLAADDGRLLYSDFLT 243

Qy 241 FLDRDDGAKKVFYVETREPRLTAAHLLFVAP--HNDS--ATGEPEASGSGPPSGGA 298

Db 244 FLDRMSSRKLFFVETROPARLLTAAHLLFVAPQHNSQSEATG-----STSG----- 292

Qy 299 LGPRALFASRVPRGQVYVVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILNR 358

Db 293 ---QALFASNVKPGQVYVLGE--GGQQLLPASVHSVLSREASGAYAPLTAQGTILNR 347

Qy 359 VLASCYAVIEHSHWAHRAFAFAPFLAHALLAALAPARTDRGDSGGGDRGGGGRVALTAP 418

Db 348 VLASCYAVIEHSHWAHRAFAFAPFLAQLLAAL-----CP 381

Qy 419 GAADAPCAGATAGIHWSYQLLYOIGTWLSDSEALHPLGMVAVKSS 462

Db 382 DGAIPTAATTTGTHWYSRLLYRIGSWLVDGDLHPLGMVAVPAS 425

RESULT 49

US-09-639-695-8

; Sequence 8, Application US/09639695

; Patent No. 6576237

; GENERAL INFORMATION:

; APPLICANT: Ingham, Phillip W.

; McMahon, Andrew P.

; Tabin, Clifford J.

; Bumcrot, David A.

; Marti-Gorostiza, Elisa

; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

; Proteins and Uses Related Thereto

; NUMBER OF SEQUENCES: 62

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/639,695

; FILING DATE: 16-Aug-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/435,093

; FILING DATE: 4-MAY-1995

; APPLICATION NUMBER: US 08/356,060

; FILING DATE: 14-DEC-1994

; APPLICATION NUMBER: US 08/176,427

; FILING DATE: 30-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: HMV-006.05

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 832-1000

; TELEFAX: (617) 832-7000

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 425 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-639-695-8

Query Match 73.2%; Score 1807; DB 4; Length 425;

Best Local Similarity 77.8%; Pred. No. 2.2e-172;

Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

Qy 1 MLLARCLLVLSLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 60

Db 4 MLLTRILLVGFICALLVSSGLTCGPGRGKRRHPKLTPLAYKQIPNVAEKTIGASG 63

SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-448-188-8

Query Match 73.2%; Score 1807; DB 4; Length 425;  
Best Local Similarity 77.8%; Pred. No. 2.2e-172;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 1 MLLARCLLLVLLVCSGLACPGPGFGRKRRHPKLTPLAYKQIPNVAEKTIGASG 60  
DB 4 MLLLRILLVGFICALLVSSGLTCGPGRGIGKRRHPKLTPLAYKQIPNVAEKTIGASG 63  
QY 61 RYEGKISNSRFRKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 120  
DB 64 RYEGKLTNSRFRKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 123  
QY 121 KLRVTEGWDEGDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
DB 124 KLRVTEGWDEGDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 183  
QY 181 IHCSVKAENSVAAGSGGCGFPGSATVHLEOGGTLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 184 IHCSVKAENSVAAGSGGCGFPGSATVHLEGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 243  
QY 241 FLDRDDGAKKVFYVITETPREPRLLLTAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA 298  
DB 244 FLDRMDSRKLFFYVITETPREPRLLLTAHLLFVAPQHNQSEATG-----STSG----- 292  
QY 299 LGPRALFASVRPGQRYVVAERDGRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINR 358  
DB 293 ---QALFASNKPGQRYVVLGE--GGQQLLPASVHVSLSREASGAYAPLTAQGTILINR 347  
QY 359 VLASCVAVIEHSHWAHRAFPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAP 418  
DB 348 VLASCVAVIEHSHWAHRAFPFRLAQLLAL-----CP 381  
QY 419 GAADAPAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462  
DB 382 DGAIPTAATTTTGIHWYSRLYRIGSVWLDGDLHPLGMVAPAS 425

Search completed: March 29, 2004, 18:34:38  
Job time : 27 secs

61 RYEGKISNSRFRKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 120  
64 RYEGKLTNSRFRKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALAISVMNQPGV 123  
121 KLRVTEGWDEGDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
124 KLRVTEGWDEGDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 183  
181 IHCSVKAENSVAAGSGGCGFPGSATVHLEOGGTLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
184 IHCSVKAENSVAAGSGGCGFPGSATVHLEGGTKLVKOLSPGDRVLAADDQGRLLYSDFLT 243  
241 FLDRDDGAKKVFYVITETPREPRLLLTAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA 298  
244 FLDRMDSRKLFFYVITETPREPRLLLTAHLLFVAPQHNQSEATG-----STSG----- 292  
299 LGPRALFASVRPGQRYVVAERDGRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINR 358  
293 ---QALFASNKPGQRYVVLGE--GGQQLLPASVHVSLSREASGAYAPLTAQGTILINR 347  
359 VLASCVAVIEHSHWAHRAFPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAP 418  
348 VLASCVAVIEHSHWAHRAFPFRLAQLLAL-----CP 381  
419 GAADAPAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462  
382 DGAIPTAATTTTGIHWYSRLYRIGSVWLDGDLHPLGMVAPAS 425

RESULT 50  
US-09-448-188-8  
Sequence 8, Application US/09448188  
Patent No. 6607913  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
McMahon, Andrew P.  
Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/448,188  
FILING DATE: 23-NOV-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 8:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 18:32:09 ; Search time 46 Seconds  
(without alignments) 2701.943 Million cell

Title: US-09-883-848A-15  
Perfect score: 2469  
Sequence: 1 MLLARCLLLVLVSS

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

```

Database : Published Applications AA:**
1: /cgn2_6/pdatat1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/pdatat1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/pdatat1/pubpaa/US05_NEW_PUB.pcp.*
4: /cgn2_6/pdatat1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/pdatat1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/pdatat1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/pdatat1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/pdatat1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/pdatat1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/pdatat1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/pdatat1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/pdatat1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/pdatat1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/pdatat1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/pdatat1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/pdatat1/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/pdatat1/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/pdatat1/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2467	99.9	475	8	US-08-900-220C-15	Sequence 15, Appl
2	2467	99.9	475	8	US-08-954-771-13	Sequence 13, Appl
3	2467	99.9	475	8	US-08-462-386D-13	Sequence 13, Appl
4	2467	99.9	475	9	US-09-021-660A-39	Sequence 39, Appl
5	2467	99.9	475	9	US-09-151-999-15	Sequence 15, Appl
6	2467	99.9	475	10	US-09-883-848A-15	Sequence 15, Appl
7	2467	99.9	475	10	US-09-187-387-15	Sequence 15, Appl
8	2467	99.9	475	10	US-09-827-110-15	Sequence 15, Appl
9	2467	99.9	475	10	US-09-845-025C-15	Sequence 15, Appl
10	2467	99.9	475	10	US-09-451-933-15	Sequence 15, Appl
11	2467	99.9	475	10	US-09-238-243-15	Sequence 15, Appl
12	2467	99.9	475	10	US-09-736-476-13	Sequence 13, Appl
13	2467	99.9	475	12	US-10-244-095A-15	Sequence 15, Appl
14	2407	97.5	462	10	US-09-733-633-14	Sequence 14, Appl
15	2407	97.5	462	15	US-10-294-036-15	Sequence 15, Appl

```

89 1271 51.5 411 10 US-09-883-848A-16
90 1271 51.5 411 10 US-09-187-387-16
91 1271 51.5 411 10 US-09-827-110-16
92 1271 51.5 411 10 US-09-845-025C-16
93 1271 51.5 411 10 US-09-451-939-16
94 1271 51.5 411 10 US-09-238-243-16
95 1271 51.5 411 12 US-10-244-095A-16
96 1266 51.3 411 10 US-09-733-634-28
97 1266 51.3 411 15 US-10-294-036-16
98 1262 51.1 449 9 US-09-990-046-29
99 1262 51.1 449 10 US-09-733-634-30
100 1248 50.5 411 15 US-10-294-036-12
101 1234 50.0 411 8 US-08-954-771-10
102 1234 50.0 411 8 US-08-462-386D-10
103 1234 50.0 411 9 US-09-151-999-12
104 1234 50.0 411 10 US-09-883-848A-12
105 1234 50.0 411 10 US-09-187-387-12
106 1234 50.0 411 10 US-09-827-110-12
107 1234 50.0 411 10 US-09-845-025C-12
108 1234 50.0 411 10 US-09-451-939-12
109 1234 50.0 411 10 US-09-238-243-12
110 1234 50.0 411 10 US-09-736-476-10
111 1234 50.0 411 12 US-10-244-095A-12
112 1234 50.0 411 8 US-08-900-220C-11
113 1183 47.9 396 8 US-08-954-771-9
114 1183 47.9 396 8 US-08-462-386D-9
115 1183 47.9 396 9 US-09-151-999-11
116 1183 47.9 396 9 US-09-021-660A-35
117 1183 47.9 396 9 US-09-151-999-11
118 1183 47.9 396 9 US-09-990-046-13
119 1183 47.9 396 10 US-09-733-634-22
120 1183 47.9 396 10 US-09-883-848A-11
121 1183 47.9 396 10 US-09-187-387-11
122 1183 47.9 396 10 US-09-827-110-11
123 1183 47.9 396 10 US-09-845-025C-11
124 1183 47.9 396 10 US-09-451-939-11
125 1183 47.9 396 10 US-09-238-243-11
126 1183 47.9 396 10 US-09-736-476-9
127 1183 47.9 396 12 US-10-244-095A-11
128 1183 47.9 396 15 US-10-294-036-11
129 1177 47.7 396 8 US-08-900-220C-17
130 1177 47.7 396 9 US-09-151-999-17
131 1177 47.7 396 9 US-09-883-848A-17
132 1177 47.7 396 10 US-09-187-387-17
133 1177 47.7 396 10 US-09-827-110-17
134 1177 47.7 396 10 US-09-845-025C-17
135 1177 47.7 396 10 US-09-451-939-17
136 1177 47.7 396 12 US-10-238-243-17
137 1177 47.7 396 12 US-10-244-095A-17
138 1177 47.7 396 15 US-10-294-036-17
139 1174 47.5 396 14 US-10-147-463-3
140 1170 47.4 396 9 US-09-244-466-2
141 1163.5 47.1 374 14 US-10-147-463-2
142 1055.5 42.8 336 9 US-09-021-660A-36
143 1031 41.8 221 8 US-08-900-220C-21
144 1031 41.8 221 8 US-08-954-771-40
145 1031 41.8 221 8 US-08-462-386D-40
146 1031 41.8 221 9 US-09-151-999-21
147 1031 41.8 221 10 US-09-187-387-21
148 1031 41.8 221 10 US-09-827-110-21
149 1031 41.8 221 10 US-09-845-025C-21
150 1031 41.8 221 10 US-09-451-939-21

```

## ALIGNMENTS

```

RESULT 1
US-08-900-220C-15
; Sequence 15, Application US/0900220C
; Publication No. US20020045206A1
; GENERAL INFORMATION:
; APPLICANT: Miao, Ningning

```

```

; Wang, Monica
; Mahanthappa, Nagesh K.
; Pang, Kevin
; Jin, Ping
;
; TITLE OF INVENTION: Method of Treating Dopaminergic and
; GABA-nergic Disorders
;
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: ONE POST OFFICE SQUARE
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,220C
; FILING DATE: 24-Jul-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: ONV-044.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
;
; US-08-900-220C-15
;
; Query Match 99.9%; Score 2467; DB 8; Length 475;
; Best Local Similarity 100.0%; Pred. No. 2,1e-207;
; Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 MLLARCLLLVSSLLVCSGLACGPGRGKGRHPKLTPLAYKQIPNVAEKTLLGASG 60
; DB 1 MLLARCLLLVSSLLVCSGLACGPGRGKGRHPKLTPLAYKQIPNVAEKTLLGASG 60
; QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAISSVNWQPGV 120
; DB 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAISSVNWQPGV 120
; QY 121 KLRVTEGWDEGHSEESLHYEGEAVDITTSDDRSKYGLMARLAVAGFDWVYESKAH 180
; DB 121 KLRVTEGWDEGHSEESLHYEGEAVDITTSDDRSKYGLMARLAVAGFDWVYESKAH 180
; QY 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
; DB 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
; QY 241 FLDRDDGKAKYFYVETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300
; DB 241 FLDRDDGKAKYFYVETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300
; QY 301 PRALFASVRPQGVVVAERDGGRRLLPAAVHVSVTLSSEAGAYAPLTAGTTLINRVL 360
; DB 301 PRALFASVRPQGVVVAERDGGRRLLPAAVHVSVTLSSEAGAYAPLTAGTTLINRVL 360
; QY 361 ASCYAVIEHSHWAHRAFPPLAHALLAALAPASTDRGGSGGDRGGGGRVALTAPGA 420
; DB 361 ASCYAVIEHSHWAHRAFPPLAHALLAALAPASTDRGGSGGDRGGGGRVALTAPGA 420
; QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSSXSRGAGGGAREGA 475

```

Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSSXRGAGGAREGA 475

RESULT 2

US-08-954-771-13

Sequence 13, Application US/08954771

Publication No. US20030054437A1

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tabin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

TITLE OF INVENTION: Proteins and Uses Related Thereto

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/954,771

FILING DATE: 20-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/462,386

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093

FILING DATE: 04-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,060

FILING DATE: 14-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMV-006.11

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 475 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-954-771-13

Query Match 99.9%; Score 2467; DB 8; Length 475;

Best Local Similarity 100.0%; Pred. No. 2.1e-207; Indels 0; Gaps 0;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKKLTPLAYKQFIPNVAEKTLAGS 60

Db 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKKLTPLAYKQFIPNVAEKTLAGS 60

QY 61 RYEGKISRNSRPFKLTNNYNDIIFKDEBNTGADRLMTORCKKLNALISVNWQPGV 120

Db 61 RYEGKISRNSRPFKLTNNYNDIIFKDEBNTGADRLMTORCKKLNALISVNWQPGV 120

QY 121 KLRVTGWDGDGHSESLHYEGRVADITTSRDRSKYGMRLARLAVERAGFDWVYESKAH 180

Db 121 KLRVTGWDGDGHSESLHYEGRVADITTSRDRSKYGMRLARLAVERAGFDWVYESKAH 180

QY 181 IHCSVKAENSVAKSGGCGPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

Db 181 IHCSVKAENSVAKSGGCGPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240

QY 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPSSGALG 300

Db 241 FLDRDDGAKKVFYVETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPSSGALG 300

QY 301 PRALFASRVPGQVYVVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360

Db 301 PRALFASRVPGQVYVVAERDGRRLPAAVHSVTLSSEAAAGAYAPLTAQGTILINRVL 360

QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGGSGGGDRGGGRVALTAPGA 420

Db 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGGSGGGDRGGGRVALTAPGA 420

QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSSXRGAGGAREGA 475

Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAKSSXRGAGGAREGA 475

RESULT 3

US-08-462-386D-13

Sequence 13, Application US/08462386D

Publication No. US20030186357A1

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tabin, Clifford J.

TITLE OF INVENTION: Vertebrate Tissue Pattern-Inducing

TITLE OF INVENTION: Proteins and Uses Related Thereto

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,386D

FILING DATE: 5-JUNE-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093

FILING DATE: 4-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,060

FILING DATE: 14-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMI-006CP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 475 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-462-386D-13

Query Match 99.9%; Score 2467; DB 8; Length 475;

Best Local Similarity 100.0%; Pred. No. 2.1e-207; Indels 0; Gaps 0;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSRERFKELTPNNPDIIFKDEBNTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
Db 61 RYEGKISRNSRERFKELTPNNPDIIFKDEBNTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
QY 121 KLRTVEGWDEBHGHSSESLHYEGRAVDITTSRDRSKYGMARLAVAEAGFDWVYVESKAH 180  
Db 121 KLRTVEGWDEBHGHSSESLHYEGRAVDITTSRDRSKYGMARLAVAEAGFDWVYVESKAH 180  
QY 181 IHC SVKAENSVAAKSGCGFPQSATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
Db 181 IHC SVKAENSVAAKSGCGFPQSATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPGQVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPGQVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
Db 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475

## RESULT 4

US-09-021-660A-39  
; Sequence 39, Application US/09021660A  
; Patent No. US20010041668A1  
; GENERAL INFORMATION:  
; APPLICANT: Baron, M.  
; APPLICANT: Farrington, S.  
; APPLICANT: Belausoff, M.  
; TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR  
; FILE REFERENCE: HUIP-P01-060  
; CURRENT APPLICATION NUMBER: US/09/021,660A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/037,513  
; PRIOR FILING DATE: 1997-02-10  
; PRIOR APPLICATION NUMBER: 60/049,763  
; PRIOR FILING DATE: 1997-06-16  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (463)  
; OTHER INFORMATION: Xaa=unknown amino acid  
US-09-021-660A-39

Query Match 99.9%; Score 2467; DB 9; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSRERFKELTPNNPDIIFKDEBNTGADRLMTQRCCKLNALAIISVMNQWPGV 120

Db 61 RYEGKISRNSRERFKELTPNNPDIIFKDEBNTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
QY 121 KLRTVEGWDEBHGHSSESLHYEGRAVDITTSRDRSKYGMARLAVAEAGFDWVYVESKAH 180  
Db 121 KLRTVEGWDEBHGHSSESLHYEGRAVDITTSRDRSKYGMARLAVAEAGFDWVYVESKAH 180  
QY 181 IHC SVKAENSVAAKSGCGFPQSATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
Db 181 IHC SVKAENSVAAKSGCGFPQSATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
QY 301 PRALFASVRPGQVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPGQVYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
Db 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDDSEALHPLGMVKSXSRGAGGAREGA 475

## RESULT 5

US-09-151-999-15  
; Sequence 15, Application US/09151999  
; Patent No. US20020151460A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Elizabeth  
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE  
; FILE REFERENCE: ONV-031.02  
; CURRENT APPLICATION NUMBER: US/09/151,999  
; CURRENT FILING DATE: 1998-08-11  
; EARLIER APPLICATION NUMBER: 08/955,552  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Homo sapien Shh  
; FEATURE:  
; OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid  
US-09-151-999-15

Query Match 99.9%; Score 2467; DB 9; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSRERFKELTPNNPDIIFKDEBNTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
Db 61 RYEGKISRNSRERFKELTPNNPDIIFKDEBNTGADRLMTQRCCKLNALAIISVMNQWPGV 120  
QY 121 KLRTVEGWDEBHGHSSESLHYEGRAVDITTSRDRSKYGMARLAVAEAGFDWVYVESKAH 180  
Db 121 KLRTVEGWDEBHGHSSESLHYEGRAVDITTSRDRSKYGMARLAVAEAGFDWVYVESKAH 180  
QY 181 IHC SVKAENSVAAKSGCGFPQSATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
Db 181 IHC SVKAENSVAAKSGCGFPQSATVHLEQGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
Db 361 ASCYAVIEHSHWAHRAFPAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 6  
US-09-883-848A-15  
; Sequence 15, Application US/09883848A  
; Publication No. US20030022819A1  
; GENERAL INFORMATION:  
; APPLICANT: Ling, L  
; APPLICANT: Sanicola-Nadel, M  
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES  
; FILE REFERENCE: CIBT-P01-119  
; CURRENT APPLICATION NUMBER: US/09/883,848A  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/211,919  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent in ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (463)  
; OTHER INFORMATION: Xaa=unknown amino acid residue  
US-09-883-848A-15

Query Match 99.9%; Score 2467; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQIPNVAEKTLCASG 60  
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQIPNVAEKTLCASG 60  
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 120  
QY 121 KLRTVEGDHEDGHSSESLHYEGRAVDITTSDDRSGYGMRLARLAVEAGFDWVYYESKAH 180  
Db 121 KLRTVEGDHEDGHSSESLHYEGRAVDITTSDDRSGYGMRLARLAVEAGFDWVYYESKAH 180  
QY 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGPEASGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGPEASGSGPPSGGALG 300  
QY 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
Db 361 ASCYAVIEHSHWAHRAFPAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475

Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 7  
US-09-187-387-15  
; Sequence 15, Application US/09187387  
; Publication No. US20030083242A1  
; GENERAL INFORMATION:  
; APPLICANT: Galdes, Alphonse  
; APPLICANT: Mahantappa, Nagesh  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING OR PREVENTING  
; TITLE OF INVENTION: PERIPHERAL NEUROPATHIES  
; FILE REFERENCE: ONV-052.01  
; CURRENT APPLICATION NUMBER: US/09/187,387  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Patent in ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: human 8hh  
; FEATURE:  
; OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid  
US-09-187-387-15

Query Match 99.9%; Score 2467; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQIPNVAEKTLCASG 60  
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQIPNVAEKTLCASG 60  
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKNALAISSVMNQPGV 120  
QY 121 KLRTVEGDHEDGHSSESLHYEGRAVDITTSDDRSGYGMRLARLAVEAGFDWVYYESKAH 180  
Db 121 KLRTVEGDHEDGHSSESLHYEGRAVDITTSDDRSGYGMRLARLAVEAGFDWVYYESKAH 180  
QY 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAAGSGGCFPGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGPEASGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGPEASGSGPPSGGALG 300  
QY 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASRVPGQVYVVAERDGRLLPAAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
Db 361 ASCYAVIEHSHWAHRAFPAPFRLAHLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLDSALHPLGMVAVKSSXSRGAGGAREGA 475

RESULT 8  
US-09-827-110-15  
; Sequence 15, Application US/09827110  
; Publication No. US20030104970A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Elizabeth  
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE  
; TITLE OF INVENTION: POLYPEPTIDES, AND FORMULATIONS AND USES RELATED THEREO  
; FILE REFERENCE: ONV-031.02  
; CURRENT APPLICATION NUMBER: US/09/827,110  
; CURRENT FILING DATE: 2000-04-05

```
; PRIOR APPLICATION NUMBER: 08/955,552
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens Shh
; FEATURE:
; OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid
US-09-827-110-15

Query Match          99.9%; Score 2467; DB 10; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.1e-207; Indels 0; Gaps 0;
Matches 475; Conservative 0; Mismatches 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60

QY 61 RYEGKISRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQCKDKLNALAI SVMNQPGV 120
Db 61 RYEGKISRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQCKDKLNALAI SVMNQPGV 120

QY 121 KLRTVEGWDGDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180
Db 121 KLRTVEGWDGDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180

QY 181 IHC SVKAENSVAKSGCGFPGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 240
Db 181 IHC SVKAENSVAKSGCGFPGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 240

QY 241 FLDRDDGAKKVFYVETREPRERLLTTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
Db 241 FLDRDDGAKKVFYVETREPRERLLTTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSBEAAGAYAPLTAQGTILINRVL 360
Db 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSBEAAGAYAPLTAQGTILINRVL 360

QY 361 ASCVAVTEEHSWAHRAFPRLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGA 420
Db 361 ASCVAVTEEHSWAHRAFPRLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGA 420

QY 421 ADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVKSXSRGAGGAREGA 475
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVKSXSRGAGGAREGA 475

RESULT 9
US-09-845-025C-15
; Sequence 15, Application US/09845025C
; Publication No. US20030104995A1
; GENERAL INFORMATION:
; APPLICANT: Reilly, J.
; TITLE OF INVENTION: NEUROPROTECTIVE METHODS AND COMPOSITIONS
; FILE REFERENCE: CIST-P01-098
; CURRENT APPLICATION NUMBER: US/09/845,025C
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/200,765
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (463)
; OTHER INFORMATION: Xaa=unknown amino acid residue
US-09-845-025C-15
```

```
Query Match          99.9%; Score 2467; DB 10; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.1e-207; Indels 0; Gaps 0;
Matches 475; Conservative 0; Mismatches 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
Db 1 MLLARCLLLVLSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60

QY 61 RYEGKISRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQCKDKLNALAI SVMNQPGV 120
Db 61 RYEGKISRNSERPKELTPNPNPDIIFKDEBNTGADRLMTQCKDKLNALAI SVMNQPGV 120

QY 121 KLRTVEGWDGDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180
Db 121 KLRTVEGWDGDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180

QY 181 IHC SVKAENSVAKSGCGFPGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 240
Db 181 IHC SVKAENSVAKSGCGFPGSATVHLEQGKTLVKDLSFGDRVLAADDOGRLLYSDFLT 240

QY 241 FLDRDDGAKKVFYVETREPRERLLTTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
Db 241 FLDRDDGAKKVFYVETREPRERLLTTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSBEAAGAYAPLTAQGTILINRVL 360
Db 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSBEAAGAYAPLTAQGTILINRVL 360

QY 361 ASCVAVTEEHSWAHRAFPRLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGA 420
Db 361 ASCVAVTEEHSWAHRAFPRLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGA 420

QY 421 ADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVKSXSRGAGGAREGA 475
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVKSXSRGAGGAREGA 475

RESULT 10
US-09-451-939-15
; Sequence 15, Application US/09451939
; Publication No. US20030119729A1
; GENERAL INFORMATION:
; APPLICANT: Miao, Ningning
; APPLICANT: Wang, Monica
; APPLICANT: Mahanthappa, Nagesh K.
; APPLICANT: Jin, Ping
; APPLICANT: Pang, Kevin
; TITLE OF INVENTION: Method of Treating Dopaminergic and
; TITLE OF INVENTION: GABA-ergic Disorders
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: ONE POST OFFICE SQUARE
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/451,939
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/900,220
; FILING DATE: 24-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: ONV-044.01
```

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-451-939-15

Query Match 99.9%; Score 2467; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTILGASG 60  
DB 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTILGASG 60  
QY 61 RYEGKISRSERFKELTPTNYPDIIFKDEENTGADRLMTQRCCKKLNALAI SVMNQPGV 120  
DB 61 RYEGKISRSERFKELTPTNYPDIIFKDEENTGADRLMTQRCCKKLNALAI SVMNQPGV 120  
QY 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
DB 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
QY 181 IHCSVKAENSVAAGSGGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAGSGGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGGSGGPGGALG 300  
DB 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGGSGGPGGALG 300  
QY 301 PRALFASVRPQGVVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASVRPQGVVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGAREGA 475

## RESULT 11

US-09-238-243-15  
Sequence 15, Application US/09238243  
Publication No. US20030162698A1  
GENERAL INFORMATION:  
APPLICANT: Galdes, Alphonse  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING DOPAMINERGIC  
TITLE OF INVENTION: AND GABA-NERGIC DISORDERS  
FILE REFERENCE: ONV-069.01  
CURRENT APPLICATION NUMBER: US/09/238,243  
CURRENT FILING DATE: 1999-01-27  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Homo sapien Shh  
FEATURE:  
OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid  
US-09-238-243-15

Query Match 99.9%; Score 2467; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTILGASG 60  
DB 1 MLLARCLLLVLSLLVCSGLACGPGRGKRRHPKLTPLAYKQIPNVAEKTILGASG 60  
QY 61 RYEGKISRSERFKELTPTNYPDIIFKDEENTGADRLMTQRCCKKLNALAI SVMNQPGV 120  
DB 61 RYEGKISRSERFKELTPTNYPDIIFKDEENTGADRLMTQRCCKKLNALAI SVMNQPGV 120  
QY 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
DB 121 KLRTVEGWDEDEGHSEESLHYEGRAVDITTSDRDRSKYGMRLARLAVEAGFDWVYYSKAH 180  
QY 181 IHCSVKAENSVAAGSGGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAGSGGCPGSA TVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGGSGGPGGALG 300  
DB 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGGSGGPGGALG 300  
QY 301 PRALFASVRPQGVVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
DB 301 PRALFASVRPQGVVVAERDGRRLIPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
DB 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWYSQLLYQIGTWLLDSEALHPLGMVAVKSSXSRGAGGAREGA 475

## RESULT 12

US-09-736-476-13  
Sequence 13, Application US/09736476  
Publication No. US20030190696A1  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
McMahon, Andrew P.  
Tabin, Clifford J.  
Bumcrot, David A.  
Martí-Gorostiza, Elisa  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/736,476  
FILING DATE: 13-Dec-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709

```
/
/ REFERENCE/DOCKET NUMBER: HMI-006CP4
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 227-7400
/ TELEFAX: (617) 227-5941
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 475 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-736-476-13

Query Match          99.9%; Score 2467; DB 10; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.1e-207;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSRERKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
DB 61 RYEGKISRNSRERKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
QY 121 KLRTVGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180
DB 121 KLRTVGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180
QY 181 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGTVLKDLSFGDRVLAADQGRLLYSDFLT 240
DB 181 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGTVLKDLSFGDRVLAADQGRLLYSDFLT 240
QY 241 FLDRDDGAKKVFVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
DB 241 FLDRDDGAKKVFVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSBEAAGAYAPLTAQGTILINRVL 360
DB 301 PRALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSBEAAGAYAPLTAQGTILINRVL 360
QY 361 ASCVAVTEESWAHRAFPRLAHALLAALAPARTDRGGDSGGDRGGGGRVALTAPGA 420
DB 361 ASCVAVTEESWAHRAFPRLAHALLAALAPARTDRGGDSGGDRGGGGRVALTAPGA 420
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSSXSRGAGGGAREGA 475
DB 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSSXSRGAGGGAREGA 475

RESULT 13
US-10-244-095A-15
; Sequence 15, Application US/10244095A
; Publication No. US20040038876A1
; GENERAL INFORMATION:
; APPLICANT: Pepinsky, Blake R.
; APPLICANT: Taylor, Frederick
; APPLICANT: Garber, Ellen A.
; FILE OF INVENTION: POLYMER CONJUGATES OF HEDGEHOG PROTEINS AND USES
; TITLE REFERENCE: CIBT-P01-117
; CURRENT FILING DATE: 2002-09-12
; PRIOR FILING DATE: 2002-09-12
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 1999-08-13
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 475
; TYPE: PRT
```

```
/
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (463)
/ OTHER INFORMATION: Xaa=unknown amino acid residue
US-10-244-095A-15

Query Match          99.9%; Score 2467; DB 12; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.1e-207;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSRERKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
DB 61 RYEGKISRNSRERKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
QY 121 KLRTVGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180
DB 121 KLRTVGWDEDDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180
QY 181 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGTVLKDLSFGDRVLAADQGRLLYSDFLT 240
DB 181 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGTVLKDLSFGDRVLAADQGRLLYSDFLT 240
QY 241 FLDRDDGAKKVFVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
DB 241 FLDRDDGAKKVFVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSBEAAGAYAPLTAQGTILINRVL 360
DB 301 PRALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSBEAAGAYAPLTAQGTILINRVL 360
QY 361 ASCVAVTEESWAHRAFPRLAHALLAALAPARTDRGGDSGGDRGGGGRVALTAPGA 420
DB 361 ASCVAVTEESWAHRAFPRLAHALLAALAPARTDRGGDSGGDRGGGGRVALTAPGA 420
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSSXSRGAGGGAREGA 475
DB 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSSXSRGAGGGAREGA 475

RESULT 14
US-09-733-634-14
; Sequence 14, Application US/09733634
; Publication No. US20030013646A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Method to stimulate Insulin production by pancreatic b-cells
; FILE REFERENCE: 17633/1240
; CURRENT APPLICATION NUMBER: US/09/733,634
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/170,282
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-733-634-14

Query Match          97.5%; Score 2407; DB 10; Length 462;
Best Local Similarity 100.0%; Pred. No. 3.7e-202;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
DB 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
QY 61 RYEGKISRNSRERKELTPNYPNDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
```

Db 61 RYEGKISRNSERFKELTPNPDIIIFKDEBENTGADRLMTQCKDKLNALAISSVNNQPGV 120  
Qy 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Db 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Qy 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Qy 301 PRALFASVRPGQVRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPGQVRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
Db 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
Qy 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
RESULT 15  
US-10-294-036-15  
; Sequence 15, Application US/10294036  
; Publication No. US2003020244A1  
; GENERAL INFORMATION:  
; APPLICANT: Warzecha, Joerg  
; TITLE OF INVENTION: HDGHHOG SIGNALING PROMOTES THE FORMATION OF THREE DIMENSIONAL CA  
; FILE REFERENCE: CIBT-P01-123  
; CURRENT APPLICATION NUMBER: US/10/294,036  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: 60/350,594  
; PRIOR FILING DATE: 2001-11-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-294-036-15  
Query Match 97.5%; Score 2407; DB 15; Length 462;  
Best Local Similarity 100.0%; Pred. No. 3.7e-202; Mismatches 0; Indels 0; Gaps 0;  
Matches 462; Conservative 0;  
Qy 1 MLLARCLLLVLVSSLLVCSGLACGPGRGFKRRHPKLLTPLAYKQFIPNVAEKTLAGS 60  
Db 1 MLLARCLLLVLVSSLLVCSGLACGPGRGFKRRHPKLLTPLAYKQFIPNVAEKTLAGS 60  
Qy 61 RYEGKISRNSERFKELTPNPDIIIFKDEBENTGADRLMTQCKDKLNALAISSVNNQPGV 120  
Db 61 RYEGKISRNSERFKELTPNPDIIIFKDEBENTGADRLMTQCKDKLNALAISSVNNQPGV 120  
Qy 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Db 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Qy 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300

Qy 301 PRALFASVRPGQVRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 301 PRALFASVRPGQVRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
Db 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
Qy 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
Db 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
RESULT 16  
US-09-021-660A-37  
; Sequence 37, Application US/09021660A  
; Patent No. US20010041688A1  
; GENERAL INFORMATION:  
; APPLICANT: Baron, M.  
; APPLICANT: Farrington, S.  
; APPLICANT: Belaussoff, M.  
; TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR  
; FILE REFERENCE: HUIP-201-060  
; CURRENT APPLICATION NUMBER: US/09/021,660A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/037,513  
; PRIOR FILING DATE: 1997-02-10  
; PRIOR APPLICATION NUMBER: 60/049,763  
; PRIOR FILING DATE: 1997-06-16  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 37  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-021-660A-37  
Query Match 83.8%; Score 2068; DB 9; Length 437;  
Best Local Similarity 87.4%; Pred. No. 1.7e-172; Mismatches 24; Indels 26; Gaps 3;  
Matches 404; Conservative 8;  
Qy 1 MLLARCLLLVLVSSLLVCSGLACGPGRGFKRRHPKLLTPLAYKQFIPNVAEKTLAGS 60  
Db 2 LLLARCLFLVLVSSLLVCSGLACGPGRGFKRRHPKLLTPLAYKQFIPNVAEKTLAGS 61  
Qy 61 RYEGKISRNSERFKELTPNPDIIIFKDEBENTGADRLMTQCKDKLNALAISSVNNQPGV 120  
Db 62 RYEGKISRNSERFKELTPNPDIIIFKDEBENTGADRLMTQCKDKLNALAISSVNNQPGV 121  
Qy 121 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180  
Db 122 KLRTVEGWDEDDHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 181  
Qy 181 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 242 FLDRDDGAKKVFYVETLEPRELLTAAHLLFVAPHND-----SQPTPG 286  
Qy 301 PRALFASVRPGQVRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPGQVRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGDRGGGGRVALTAPGA 420  
Db 347 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPART-----GGGGGSIIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAVKSS 462  
Db 396 TEARGAEPAGIHWSQLLYHIGTWLLDSETWHPLGMVAVKSS 437

```
RESULT 17
US-09-969-520A-1
; Sequence 1, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JH01670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-969-520A-1

Query Match      83.8%; Score 2068; DB 9; Length 437;
Best Local Similarity 87.4%; Pred. No. 1.7e-172;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
DB 2 LLLARCFVLVILASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 61
QY 61 RYEGKISRNSERPKELTPNNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQWPGV 120
DB 62 RYEGKITRNSERPKELTPNNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQWPGV 121
QY 121 KLRVTEGWDEDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180
DB 122 KLRVTEGWDEDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 181
QY 181 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 182 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDCAKVKFVVIETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
DB 242 FLDRDEGAKVKFVVIETLEPRELLTAAHLLFVAPHND-----SGPTPG 286
QY 301 PRALFASRVPGQGVVVAERDGRLLPAAVHSVTLSEAGAYAPLTAQGTILINRVL 360
DB 287 PSALFASRVPGQGVVVAERDGRLLPAAVHSVTLSEAGAYAPLTAHGTILINRVL 346
QY 361 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTDRGSDGGGGRVALTAPGA 420
DB 347 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395

RESULT 18
US-09-733-634-16
; Sequence 16, Application US/09733634
; Publication No. US20030003646A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Method to stimulate insulin production by pancreatic b-cells
; FILE REFERENCE: 17633/1240
; CURRENT APPLICATION NUMBER: US/09/733,634
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/170,282
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: House Mouse
US-10-013-310-1

Query Match      83.8%; Score 2068; DB 13; Length 437;
Best Local Similarity 87.4%; Pred. No. 1.7e-172;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-634-16

Query Match      83.8%; Score 2068; DB 10; Length 437;
Best Local Similarity 87.4%; Pred. No. 1.7e-172;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 60
DB 2 LLLARCFVLVILASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTIGASG 61
QY 61 RYEGKISRNSERPKELTPNNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQWPGV 120
DB 62 RYEGKITRNSERPKELTPNNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQWPGV 121
QY 121 KLRVTEGWDEDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180
DB 122 KLRVTEGWDEDGHSESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 181
QY 181 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 182 IHCSVKAENSVAAKSGCGPFGSATVHLEOGGKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDCAKVKFVVIETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
DB 242 FLDRDEGAKVKFVVIETLEPRELLTAAHLLFVAPHND-----SGPTPG 286
QY 301 PRALFASRVPGQGVVVAERDGRLLPAAVHSVTLSEAGAYAPLTAQGTILINRVL 360
DB 287 PSALFASRVPGQGVVVAERDGRLLPAAVHSVTLSEAGAYAPLTAHGTILINRVL 346
QY 361 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTDRGSDGGGGRVALTAPGA 420
DB 347 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395

RESULT 19
US-10-013-310-1
; Sequence 1, Application US/10013310
; Publication No. US20020192216A1
; GENERAL INFORMATION:
; APPLICANT: Lamb, Jonathon Robert
; APPLICANT: Hoyne, Gerard Francis
; APPLICANT: Dallman, Margaret Jane
; TITLE OF INVENTION: Therapeutic Use
; FILE REFERENCE: 674525-2003
; CURRENT APPLICATION NUMBER: US/10/013,310
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/GB00/02191
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: UK 9913350.6
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: UK 9921953.7
; PRIOR FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: House Mouse
US-10-013-310-1

Query Match      83.8%; Score 2068; DB 13; Length 437;
Best Local Similarity 87.4%; Pred. No. 1.7e-172;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;
```

```

QY 1 MLLARCLLVVSLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 2 LLLARCLLVVSLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKNALAISVMNQWPGV 120
DB 62 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKNALAISVMNQWPGV 121
QY 121 KLRVTEGDEHGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYESKAH 180
DB 122 KLRVTEGDEHGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYESKAH 181
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDGAKKVYVETLEPRERLLTAAHLLFVAPHNDSATGEPEASGSGPPSGGALG 300
DB 242 FLDRDGAKKVYVETLEPRERLLTAAHLLFVAPHND-----SGTPG 286
QY 301 PRALFASVRPGQVYVVAERDGRRLPAAVHVTLSSEAGAYAPLTAOGTILINRVL 360
DB 287 PSALFASVRPGQVYVVAERDGRRLPAAVHVTLSSEAGAYAPLTAOGTILINRVL 346
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420
DB 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395
QY 421 ADAPGAGTAGIHWYSQLLYQIGTWLDDSEALHPLGMVAVKSS 462
DB 396 TEARGAETAGIHWYSQLLYHIGTWLDDSETMHPGLGMVAVKSS 437

```

## RESULT 20

```

US-10-460-594-20
; Sequence 20, Application US/10460594
; Publication No. US20040018979A1
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/460,594
; FILING DATE: 11-Jun-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/061,323
; FILING DATE: 07-OCT-1997
; APPLICATION NUMBER: 08/729,743
; FILING DATE: 07-OCT-1996
; APPLICATION NUMBER: 08/567,357
; FILING DATE: 10-JUL-1996
; APPLICATION NUMBER: 08/349,498
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: <Unknown>
; FILING DATE: 02-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/140001

```

## TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-460-594-20

```

Query Match 83.8%; Score 2068; DB 15; Length 437;

Best Local Similarity 87.4%; Pred. No. 1.7e-172;

Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

```

QY 1 MLLARCLLVVSLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 2 LLLARCLLVVSLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKNALAISVMNQWPGV 120
DB 62 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKNALAISVMNQWPGV 121
QY 121 KLRVTEGDEHGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYESKAH 180
DB 122 KLRVTEGDEHGHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYYESKAH 181
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
DB 182 IHCSVKAENSVAAGSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDGAKKVYVETLEPRERLLTAAHLLFVAPHNDSATGEPEASGSGPPSGGALG 300
DB 242 FLDRDGAKKVYVETLEPRERLLTAAHLLFVAPHND-----SGTPG 286
QY 301 PRALFASVRPGQVYVVAERDGRRLPAAVHVTLSSEAGAYAPLTAOGTILINRVL 360
DB 287 PSALFASVRPGQVYVVAERDGRRLPAAVHVTLSSEAGAYAPLTAOGTILINRVL 346
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420
DB 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395
QY 421 ADAPGAGTAGIHWYSQLLYQIGTWLDDSEALHPLGMVAVKSS 462
DB 396 TEARGAETAGIHWYSQLLYHIGTWLDDSETMHPGLGMVAVKSS 437

```

## RESULT 21

```

US-08-900-220C-13
; Sequence 13, Application US/08900220C
; Publication No. US20020045206A1
; GENERAL INFORMATION:
; APPLICANT: Miao, Ningning
; Wang, Monica
; Mahanthappa, Nagesh K.
; Pang, Kevin
; Jin, Ping
; TITLE OF INVENTION: Method of Treating Dopaminergic and
; GABA-nergic Disorders
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: ONE POST OFFICE SQUARE
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
,
,
, SOFTWARE: Ascii (text)
, CURRENT APPLICATION DATA:
,   APPLICATION NUMBER: US/08/900,220C
,   FILING DATE: 24-Jul-1997
,   CLASSIFICATION: <Unknown>
, ATTORNEY/AGENT INFORMATION:
,   NAME: Vincent, Matthew P.
,   REGISTRATION NUMBER: 36,709
,   REFERENCE/DOCKET NUMBER: ONV-044.01
, TELECOMMUNICATION INFORMATION:
,   TELEPHONE: (617) 832-1000
,   TELEFAX: (617) 832-7000
, INFORMATION FOR SEQ ID NO: 13:
, SEQUENCE CHARACTERISTICS:
,   LENGTH: 437 amino acids
,   TYPE: amino acid
,   TOPOLOGY: linear
, MOLECULE TYPE: protein
, SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-08-900-220C-13

Query Match      83.6%; Score 2065; DB 8; Length 437;
Best Local Similarity 87.2%; Pred. No. 3.2e-172;
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
Db 2 LLLARCLFVLASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISVMNQPGV 120
Db 62 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISVMNQPGV 121
QY 121 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
Db 122 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181
QY 181 IHCSVKAENSVAKSGGCGFGSATVHLEQGTGLVKDLSPGDRVLAADDQGRLLYSDFLT 240
Db 182 IHCSVKAENSVAKSGGCGFGSATVHLEQGTGLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
Db 242 FLDRDEGAKKVFYVETLEPRERILLTAHLLFVAPHND-----SGPTPG 286
QY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHVSVTLSSEEAAGAYAPLTAQGTILINRVL 360
Db 287 PSALFASRVPRGQRYVVAERDGRRLPAAVHVSVTLSSEEAAGAYAPLTAHGTILINRVL 346
QY 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
Db 347 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWYSOLLVYQIGTWLDDSEALHPLGMVAKSS 462
Db 396 TEARGAETAGIHWYSOLLVYHIGTWLDDSETMHPHGMVAKSS 437

RESULT 22
US-08-954-771-11
, Sequence 11, Application No./08954771
, Publication No. US20030054437A1
, GENERAL INFORMATION:
, APPLICANT: Ingham, Phillip W.
, APPLICANT: McMahon, Andrew P.
, APPLICANT: Tabin, Clifford J.
, TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
, TITLE OF INVENTION: Proteins and Uses Related Thereto
, NUMBER OF SEQUENCES: 48
, CORRESPONDENCE ADDRESS:
, ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
, STREET: One Post Office Square
, CITY: Boston
```

```
,
,
, STATE: MA
, COUNTRY: USA
, ZIP: 02109-2170
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
,   APPLICATION NUMBER: US/08/954,771
,   FILING DATE: 20-OCT-1997
,   PRIOR APPLICATION DATA:
,     APPLICATION NUMBER: US 08/462,386
,     FILING DATE: 05-JUN-1995
,   PRIOR APPLICATION DATA:
,     APPLICATION NUMBER: US 08/435,093
,     FILING DATE: 04-MAY-1995
,   PRIOR APPLICATION DATA:
,     APPLICATION NUMBER: US 08/356,060
,     FILING DATE: 14-DEC-1994
,   PRIOR APPLICATION DATA:
,     APPLICATION NUMBER: US 08/176,427
,     FILING DATE: 30-DEC-1993
, ATTORNEY/AGENT INFORMATION:
,   NAME: Vincent, Matthew P.
,   REGISTRATION NUMBER: 36,709
,   REFERENCE/DOCKET NUMBER: HMV-006.11
, TELECOMMUNICATION INFORMATION:
,   TELEPHONE: 617-832-1000
,   TELEFAX: 617-832-7000
, INFORMATION FOR SEQ ID NO: 11:
, SEQUENCE CHARACTERISTICS:
,   LENGTH: 437 amino acids
,   TYPE: amino acid
,   TOPOLOGY: linear
, MOLECULE TYPE: protein
, US-08-954-771-11

Query Match      83.6%; Score 2065; DB 8; Length 437;
Best Local Similarity 87.2%; Pred. No. 3.2e-172;
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
Db 2 LLLARCLFVLASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISVMNQPGV 120
Db 62 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTQCKDKLNALAISVMNQPGV 121
QY 121 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
Db 122 KLRTVEGDWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181
QY 181 IHCSVKAENSVAKSGGCGFGSATVHLEQGTGLVKDLSPGDRVLAADDQGRLLYSDFLT 240
Db 182 IHCSVKAENSVAKSGGCGFGSATVHLEQGTGLVKDLSPGDRVLAADDQGRLLYSDFLT 241
QY 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
Db 242 FLDRDEGAKKVFYVETLEPRERILLTAHLLFVAPHND-----SGPTPG 286
QY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHVSVTLSSEEAAGAYAPLTAQGTILINRVL 360
Db 287 PSALFASRVPRGQRYVVAERDGRRLPAAVHVSVTLSSEEAAGAYAPLTAHGTILINRVL 346
QY 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDRGDSGGDRGGGGRVALTAPGA 420
Db 347 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
QY 421 ADAPGAGATAGIHWYSOLLVYQIGTWLDDSEALHPLGMVAKSS 462
Db 396 TEARGAETAGIHWYSOLLVYHIGTWLDDSETMHPHGMVAKSS 437
```



## RESULT 25

US-09-883-848a-13  
; Sequence 13, Application US/09883848A  
; Publication No. US20030022819A1  
; GENERAL INFORMATION:

; APPLICANT: Ling, L.  
; APPLICANT: Saricola-Nadel, M.  
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES

; FILE REFERENCE: CIBT-P01-119

; CURRENT APPLICATION NUMBER: US/09/883,848A

; CURRENT FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: 60/211,919

; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 437

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-883-848a-13

Query Match 83.6%; Score 2065; DB 10; Length 437;

Best Local Similarity 87.2%; Pred. No. 3.2e-172; Indels 26; Gaps 3;

Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTGASG 60

Db 2 LLLARCFVLIVLASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTGASG 61

QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQPGV 120

Db 62 RYEGKITRNSERFKELTPNYPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQPGV 121

QY 121 KLRVTEGWDGHHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180

Db 122 RLRVTEGWDGHHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 181

QY 181 IHC SVKAENSVAAKSGCGPGSATVHLEOGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240

Db 182 IHC SVKAENSVAAKSGCGPGSATVHLEOGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 241

QY 241 FLDRDGAQKVFVVIETREPRERLLTAAHLFPVAPHNDSATGEPEASSGSGPPSGGALG 300

Db 242 FLDRDGAQKVFVVIETREPRERLLTAAHLFPVAPHNDSATGEPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPGQVRVVAERDGRLLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 360

Db 287 PSALFASRVPGQVRVVAERDGRLLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 346

QY 361 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420

Db 347 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 395

QY 421 ADAPGAGATAGIHWYSQLLYQICTWLLDSEALPLGMVAVKSS 462

Db 396 TEARGAETAGIHWYSQLLYQICTWLLDSEALPLGMVAVKSS 437

## RESULT 26

US-09-187-387-13

; Sequence 13, Application US/09187387

; Publication No. US20030083242A1

; GENERAL INFORMATION:

; APPLICANT: Galdes, Alphonse

; APPLICANT: Mahanadap, Nagesh

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING OR PREVENTING

; FILE REFERENCE: ONV-052.01

; CURRENT APPLICATION NUMBER: US/09/187,387

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 437

; TYPE: PRT

; ORGANISM: mouse Shh

US-09-187-387-13

Query Match 83.6%; Score 2065; DB 10; Length 437;

Best Local Similarity 87.2%; Pred. No. 3.2e-172;

Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTGASG 60

Db 2 LLLARCFVLIVLASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTGASG 61

QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQPGV 120

Db 62 RYEGKITRNSERFKELTPNYPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQPGV 121

QY 121 KLRVTEGWDGHHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 180

Db 122 RLRVTEGWDGHHSEESLHYEGRAVDITTSRDRSKYGMRLARLAVEAGFDWVYYESKAH 181

QY 181 IHC SVKAENSVAAKSGCGPGSATVHLEOGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240

Db 182 IHC SVKAENSVAAKSGCGPGSATVHLEOGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 241

QY 241 FLDRDGAQKVFVVIETREPRERLLTAAHLFPVAPHNDSATGEPEASSGSGPPSGGALG 300

Db 242 FLDRDGAQKVFVVIETREPRERLLTAAHLFPVAPHNDSATGEPEASSGSGPPSGGALG 300

QY 301 PRALFASRVPGQVRVVAERDGRLLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 360

Db 287 PSALFASRVPGQVRVVAERDGRLLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 346

QY 361 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 420

Db 347 ASCVAVIEHSHWAHRAFPAPRLAHALLAALAPARTDRGDSGGDRGGGRVALTAPGA 395

QY 421 ADAPGAGATAGIHWYSQLLYQICTWLLDSEALPLGMVAVKSS 462

Db 396 TEARGAETAGIHWYSQLLYQICTWLLDSEALPLGMVAVKSS 437

## RESULT 27

US-09-827-110-13

; Sequence 13, Application US/09827110

; Publication No. US20030104970A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Elizabeth

; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE

; FILE REFERENCE: ONV-031.02

; CURRENT APPLICATION NUMBER: US/09/827,110

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: 08/955,552

; PRIOR FILING DATE: 1997-10-20

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 437

; TYPE: PRT

; ORGANISM: murine Shh

US-09-827-110-13

Query Match 83.6%; Score 2065; DB 10; Length 437;

Best Local Similarity 87.2%; Pred. No. 3.2e-172;

Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTGASG 60

Db 2 LLLARCFVLIVLASSLLVCPGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTGASG 61

QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEBNTGADRLMTQRCCKLNALAI SVMNQPGV 120

Db 62 RYEGKITRNSRPFKLTPTNPNPDIIFKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 121  
Qy 121 KLRVTEGWDGDHSEESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 180  
Db 122 RLRVTEGWDGDHSEESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 181  
Qy 181 IHCYKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCYKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETREPRERILLTAHLLFVAPHND-----SGTPTG 286  
Qy 301 PRALFASRVPRGQRYVVAERDGRRLPPAHSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASRVPRGQRYVVAERDGRRLPPAHSVTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTD-----GGGGSIP-AAQSA 420  
Db 347 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTD-----GGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWYSQLLYOIGTWLLDSEALHPLGMVAKSS 462  
Db 396 TEARGABPTAGIHWYSQLLYOIGTWLLDSEALHPLGMVAKSS 437  
RESULT 28  
US-09-845-025C-13  
; Sequence 13, Application US/09845025C  
; Publication No. US20030104995A1  
; GENERAL INFORMATION:  
; APPLICANT: Reilly, J.  
; TITLE OF INVENTION: NEUROPROTECTIVE METHODS AND COMPOSITIONS  
; FILE REFERENCE: CIBT-P01-098  
; CURRENT APPLICATION NUMBER: US/09/845,025C  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 60/200,765  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-845-025C-13  
Query Match 83.6%; Score 2065; DB 10; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.2e-172;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLFLVILASSLLVCFGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
Qy 61 RYEGKISRNSRPFKLTPTNPNPDIIFKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 120  
Db 62 RYEGKITRNSRPFKLTPTNPNPDIIFKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 121  
Qy 121 KLRVTEGWDGDHSEESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 180  
Db 122 RLRVTEGWDGDHSEESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 181  
Qy 181 IHCYKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCYKAENSVAKSGCGPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVFYVETREPRERILLTAHLLFVAPHNDSATGPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFYVETREPRERILLTAHLLFVAPHND-----SGTPTG 286  
Qy 301 PRALFASRVPRGQRYVVAERDGRRLPPAHSVTLSEEAAGAYAPLTAQGTILINRVL 360

Db 287 PSALFASRVPRGQRYVVAERDGRRLPPAHSVTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTD-----GGGGSIP-AAQSA 420  
Db 347 ASCYAVIEHSHWAHRAFPRLAHALLAALAPARTD-----GGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWYSQLLYOIGTWLLDSEALHPLGMVAKSS 462  
Db 396 TEARGABPTAGIHWYSQLLYOIGTWLLDSEALHPLGMVAKSS 437  
RESULT 29  
US-09-451-939-13  
; Sequence 13, Application US/09451939  
; Publication No. US20030119729A1  
; GENERAL INFORMATION:  
; APPLICANT: Miao, Ningning  
; APPLICANT: Wang, Monica  
; APPLICANT: Mahanthappa, Nagesh K.  
; APPLICANT: Jin, Ping  
; APPLICANT: Pang, Kevin  
; TITLE OF INVENTION: Method of Treating Dopaminergic and  
; TITLE OF INVENTION: GABA-ergic Disorders  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: ONE POST OFFICE SQUARE  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Ascii (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/451,939  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/900,220  
; FILING DATE: 24-JUL-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: ONV-044.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-451-939-13  
Query Match 83.6%; Score 2065; DB 10; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.2e-172;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
Qy 1 MLLARCLLLVLSLLVCSGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLFLVILASSLLVCFGLACGPGRGFKRRHPKLTPLAYKQFIPNVAEKTILGASG 61  
Qy 61 RYEGKISRNSRPFKLTPTNPNPDIIFKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 120  
Db 62 RYEGKITRNSRPFKLTPTNPNPDIIFKDEBNTGADRLMTQRCCKLNALAISSVMNQPGV 121  
Qy 121 KLRVTEGWDGDHSEESLHYEGRAVDITTSDDRSKYGLMARLAVAEAGFDWVYVESKAH 180

Db 122 RLRTVEGDWDEGHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFVYIETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFVYIETREPRERLLTAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASVRPQGVYVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAOGTILINRVL 360  
Db 287 PSALFASVRPQGVYVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAOGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGIP-AAQSA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVYKSS 462  
Db 396 TEARGAEPATAGIHWSQLLYQIGTWLLDSEALHPLGMVYKSS 437  
RESULT 30  
US-09-238-243-13  
; Sequence 13, Application US/09238243  
; Publication No. US20030162698A1  
; GENERAL INFORMATION:  
; APPLICANT: Galdes, Alphonse  
; APPLICANT: Mahanchappa, Nagesh  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING DOPAMINERGIC  
; FILE REFERENCE: ONV-069.01  
; CURRENT APPLICATION NUMBER: US/09/238,243  
; CURRENT FILING DATE: 1999-01-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: murine Shh  
US-09-238-243-13  
Query Match 83.6%; Score 2065; DB 10; Length 437;  
Best Local Similarity 87.2%; Pred No. 3.2e-172;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISNSRERKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALALISVWNPQGV 120  
Db 62 RYEGKITRNSRERKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALALISVWNPQGV 121  
QY 121 KLRVTEGDWDEGHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180  
Db 122 RLRTVEGDWDEGHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 181  
QY 181 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGGCPFGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVFVYIETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDEGAKKVFVYIETREPRERLLTAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASVRPQGVYVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAOGTILINRVL 360  
Db 287 PSALFASVRPQGVYVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAOGTILINRVL 346  
QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGIP-AAQSA 420

Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVYKSS 462  
Db 396 TEARGAEPATAGIHWSQLLYQIGTWLLDSEALHPLGMVYKSS 437  
RESULT 31  
US-09-736-476-11  
; Sequence 11, Application US/09736476  
; Publication No. US20030190696A1  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; McMahon, Andrew P.  
; Tabin, Clifford J.  
; Bumcrot, David A.  
; Marti-Corostiza, Elisa  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/736,476  
; FILING DATE: 13-Dec-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 4-MAY-1995  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HMI-006CP4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-736-476-11  
Query Match 83.6%; Score 2065; DB 10; Length 437;  
Best Local Similarity 87.2%; Pred No. 3.2e-172;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60  
Db 2 LLLARCLLVVSSLLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 61  
QY 61 RYEGKISNSRERKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALALISVWNPQGV 120  
Db 62 RYEGKITRNSRERKELTPNYPDIIFKDEENTGADRLMTQRCCKLNALALISVWNPQGV 121  
QY 121 KLRVTEGDWDEGHSESLHYEGRVADITTSRDRSKYGMRLARLAVEAGFDWVYVESKAH 180

Db 122 RLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 181  
Qy 181 IHCSVKAENSVAAKSGGCGPFGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAKSGGCGPFGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVYVYIETREPRERLLTAAHLLFVAPHND SATGEPEASSGGPPSGGALG 300  
Db 242 FLDRDEGAKKVYVYIETLEPRERLLTAAHLLFVAPHND -----SGTPTG 286  
Qy 301 PRALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDGGDRGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLDDSEALHPLGMVAVKSS 462  
Db 396 TEARGAEPTAGIHWYSQLLYHIGTWLDDSETMHPGLMAVKSS 437  
RESULT 32  
US-10-244-095A-13  
; Sequence 13, Application US/10244095A  
; Publication No. US20040038876A1  
; GENERAL INFORMATION:  
; APPLICANT: Pepinsky, Blake R.  
; APPLICANT: Taylor, Frederick  
; APPLICANT: Garber, Ellen A.  
; TITLE OF INVENTION: POLYMER CONJUGATES OF HEDGEHOG PROTEINS AND USES  
; FILE REFERENCE: CIBT-P01-117  
; CURRENT APPLICATION NUMBER: US/10/244,095A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: PCT/US00/14741  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/149,016  
; PRIOR FILING DATE: 1999-08-13  
; PRIOR APPLICATION NUMBER: 60/137,011  
; PRIOR FILING DATE: 1999-08-01  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-244-095A-13

Query Match 83.6%; Score 2065; DB 12; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.2e-172;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
Qy 1 MLLARCLLLVLVSSLLVCSGLACGPGRGKRRHPKKLTPLAYKQFIPNVAEKTLGASG 60  
Db 2 LLLARCFVLIVASSLLVCPGLACGPGRGKRRHPKKLTPLAYKQFIPNVAEKTLGASG 61  
Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 121  
Qy 121 KLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
Db 122 RLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 181  
Qy 181 IHCSVKAENSVAAKSGGCGPFGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAKSGGCGPFGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVYVYIETREPRERLLTAAHLLFVAPHND SATGEPEASSGGPPSGGALG 300  
Db 242 FLDRDEGAKKVYVYIETLEPRERLLTAAHLLFVAPHND -----SGTPTG 286  
Qy 301 PRALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDGGDRGGGRVALTAPGA 420  
Db 396 TEARGAEPTAGIHWYSQLLYHIGTWLDDSETMHPGLMAVKSS 437  
RESULT 34  
US-09-990-046-14

Qy 301 PRALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDGGDRGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLDDSEALHPLGMVAVKSS 462  
Db 396 TEARGAEPTAGIHWYSQLLYHIGTWLDDSETMHPGLMAVKSS 437  
RESULT 33  
US-10-294-036-13  
; Sequence 13, Application US/10294036  
; Publication No. US20030220244A1  
; GENERAL INFORMATION:  
; APPLICANT: Warzecha, Joerg  
; TITLE OF INVENTION: HEDGEHOG SIGNALING PROMOTES THE FORMATION OF THREE DIMENSIONAL C  
; FILE REFERENCE: CIBT-P01-123  
; CURRENT APPLICATION NUMBER: US/10/294,036  
; PRIOR FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: 60/350,594  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-294-036-13

Query Match 83.6%; Score 2065; DB 15; Length 437;  
Best Local Similarity 87.2%; Pred. No. 3.2e-172;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
Qy 1 MLLARCLLLVLVSSLLVCSGLACGPGRGKRRHPKKLTPLAYKQFIPNVAEKTLGASG 60  
Db 2 LLLARCFVLIVASSLLVCPGLACGPGRGKRRHPKKLTPLAYKQFIPNVAEKTLGASG 61  
Qy 61 RYEGKISRNSERFKELTPNYPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 120  
Db 62 RYEGKITRNSERFKELTPNYPDIIFKDENTGADRLMTQCKDKLNALAI SVMNQPGV 121  
Qy 121 KLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 180  
Db 122 RLRTVEGDHDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAH 181  
Qy 181 IHCSVKAENSVAAKSGGCGPFGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAKSGGCGPFGSATVHLEQGQTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241  
Qy 241 FLDRDDGAKKVYVYIETREPRERLLTAAHLLFVAPHND SATGEPEASSGGPPSGGALG 300  
Db 242 FLDRDEGAKKVYVYIETLEPRERLLTAAHLLFVAPHND -----SGTPTG 286  
Qy 301 PRALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPQGVVVAERDGRRLPAAVHVSUTLSEEAAGAYAPLTAHGTILINRVL 346  
Qy 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDGGDRGGGRVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
Qy 421 ADAPGAGATAGIHWYSQLLYQIGTWLDDSEALHPLGMVAVKSS 462  
Db 396 TEARGAEPTAGIHWYSQLLYHIGTWLDDSETMHPGLMAVKSS 437  
RESULT 34  
US-09-990-046-14

```

; Sequence 14, Application US/09990046
; Patent No. US2002015624SAI
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/990,046
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 14
; LENGTH: 437.
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-990-046-14

Query Match      83.5%; Score 2062; DB 9; Length 437;
Best Local Similarity 87.0%; Pred.No.5.8e-172;
Matches 402; Conservative 10; Mismatches 24; Indels 26; Gaps 3;

QY      1  MLLARCLLLVLVSVLLVCGSLACGPGRGFGRRHPKKLTPLAYKQFIPNVAEKTLCASG 60
DB      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      2  LLLARCFVLIVASSLLVCPGLACGPGRGFGRRHPKKLTPLAYKQFIPNVAEKTLCASG 61
DB      2  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61 RYEGKISRNSRFKELTPNNPDIIKFKEENTGADRLMTQCKOKLNLAISVNNQPGV 120
DB      61 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      62 RYEGK1TRNSRFKELTPNNPDIIKFKEENTGADRLMTQCKOKLNLAISVNNQPGV 121
DB      62 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      121 KLRVTEGWDGDGHGSESLHYEGRAVDITTSDDRSKYCMRLARLAVEAGFQWVYYESKAH 180
DB      121 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      122 KLRVTEGWDGDGHGSESLHYEGRAVDITTSDDRSKYCMRLARLAVEAGFQWVYYESKAH 181
DB      122 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      181 IHC5VKAEN5VAAKSGCGCPGSGATVHLEQCGTKLYKDI1SPGRVLAADDQGRILYSDFLT 240
DB      181 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      182 IHC5VKAEN5VAAKSGCGCPGSGATVHLEQCGTKLYKDI1SPGRVLAADDQGRILYSDFLT 241
DB      182 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      241 FLDRDDGAKKVFYV1ETREPRERLLI1TAHLLFVAPHNDSATGEPSEASSGGSPGGGALG 300
DB      241 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      242 FLDRD5GAKKVFYV1ETLEPRERLLI1TAHLLFVAPHND-----SGPTPG 286
DB      242 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      301 PRALFASRVPCQGVVYVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAOGTILINRVL 360
DB      301 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      287 PSALFASRVPCQGVVYVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAOGTILINRVL 346
DB      287 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      361 ASCYAVIEH5WAHRAFAFRLAHALLAALAPARTDRGCGDGGGGRGGGRVALTPGA 420
DB      361 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      347 ASCYAVIEH5WAHRAFAFRLAHALLAALAPARTD-----GGGGG5IP-AAQSA 395
DB      347 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      421 ADAPGAGATAGIHWISQLLYQ1GTWLLDSEALHPLGMVKSS 462
DB      421 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      396 TEARGAEP7AGIHWISQLLYH1GTWLLDSETWHPLGMVAKS 437
DB      396 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 35
US-09-969-520A-3
; Sequence 3, Application US/09969520A
; Patent No. US2002017163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JHU1670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 437
; TYPE: PRT

```

```
Db 62 RYEGKITNSERFKELTPNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVNNQPGV 121
Qy 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
Db 122 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181
Qy 181 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
Db 182 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGGPPSGGALG 300
Db 242 FLDRDEGAKKVFYVETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
Qy 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360
Db 287 PSPLFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 346
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGGGRGGRVALTAP-G 419
Db 347 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTD-----GGGG--SIPAPQS 394
Qy 420 AADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462
Db 395 VAERGAGPPAGIHWSQLLYHIGTWLSDSETMHPGLGMVAKSS 437
RESULT 37
US-09-969-520A-10
; Sequence 10, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JH01670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; PRIOR FILING DATE: 2002-05-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-10
Query Match 83.4%; Score 2059; DB 9; Length 437;
Best Local Similarity 87.0%; Pred. No. 1.1e-171;
Matches 402; Conservative 9; Mismatches 25; Indels 26; Gaps 3;
Qy 1 MILLARCLLLVLSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
Db 2 LLLARCFVLIVLASSLLVCPGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
Qy 61 RYEGKISNSERFKELTPNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVNNQPGV 120
Db 62 RYEGKITNSERFKELTPNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVNNQPGV 121
Qy 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
Db 122 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 181
Qy 181 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
Db 182 LLLARCFVLIVLASSLLVCPGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGGPPSGGALG 300
Db 242 FLDRDEGAKKVFYVETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
Qy 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360
Db 287 PSPLFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 346
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGGGRGGRVALTAPGA 420
Db 347 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
Qy 421 AADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462
Db 396 TEARGAEPAGIHWSQLLYHIGTWLSDSETMHPGLGMVAKSS 437
```

```
Db 242 FLDRDEGAKKVFYVETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
Qy 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360
Db 287 PSPLFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 346
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGGGRGGRVALTAPGA 420
Db 347 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
Qy 421 AADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462
Db 396 TEARGAEPAGIHWSQLLYHIGTWLSDSETMHPGLGMVAKSS 437
RESULT 38
US-09-969-520A-11
; Sequence 11, Application US/09969520A
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: BEACHY, Philip A.
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
; FILE REFERENCE: JH01670-1
; CURRENT APPLICATION NUMBER: US/09/969,520A
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/235,153
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Altered sonic hedgehog protein
US-09-969-520A-11
Query Match 83.4%; Score 2059; DB 9; Length 437;
Best Local Similarity 87.0%; Pred. No. 1.1e-171;
Matches 402; Conservative 9; Mismatches 25; Indels 26; Gaps 3;
Qy 1 MILLARCLLLVLSLLVCSGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
Db 2 LLLARCFVLIVLASSLLVCPGLACGPGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
Qy 61 RYEGKISNSERFKELTPNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVNNQPGV 120
Db 62 RYEGKITNSERFKELTPNYNDIIFKDEENTGADRLMTQRCCKLNALAISSVNNQPGV 121
Qy 121 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVESKAH 180
Db 122 KLRTVEGWDEGHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWVYVEAAAH 181
Qy 181 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240
Db 182 IHCSVKAENSVAAKSGGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 241
Qy 241 FLDRDDGAKKVFYVETREPRERLLTAAHLLFVAPHNDSATGPEASSGGPPSGGALG 300
Db 242 FLDRDEGAKKVFYVETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
Qy 301 PRALFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 360
Db 287 PSPLFASVRPGQRYVVAERDGRRLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVL 346
Qy 361 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTDRGSDGGGGRGGRVALTAPGA 420
Db 347 ASCYAVIEEHSWAHRAFPFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395
Qy 421 AADAPGAGATAGIHWSQLLYQIGTWLSDSEALHPLGMVAKSS 462
Db 396 TEARGAEPAGIHWSQLLYHIGTWLSDSETMHPGLGMVAKSS 437
```

RESULT 39  
US-09-969-520A-2  
; Sequence 2, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Philip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-2

Query Match 83.4%; Score 2058; DB 9; Length 437;  
Best Local Similarity 87.2%; Pred. No. 1.3e-171;  
Matches 403; Conservative 8; Mismatches 25; Indels 26; Gaps 3;

Qy 1 MILLARCLLLVSSILVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 2 LLLARCFVLVILASSLLVCFGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 61

Qy 61 RYEGKISRNSRPFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQPGV 120  
Db 62 RYEGKITNSRPFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQPGV 121

Qy 121 KLRTGWDGDDGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWYYESKAH 180  
Db 122 KLRTGWDGDDGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWYYESKAH 181

Qy 181 IHCSVKAENSVAAGSGCGFFGSAIVHLEQGTGLVKDLSGDRVLAADQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGCGFFGSAIVHLEQGTGLVKDLSGDRVLAADQGRLLYSDFLT 241

Qy 241 FLDRDGAKKVYVIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDGAKKVYVIETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286

Qy 301 PRALFASVRPGQVYVVAERDGRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPGQVYVVAERDGRLLPAAVHVSVTLSEEAAGAYAPLTAHGTILINRVL 346

Qy 361 ASCVAVIEHSHWAHRAFPAPFLAHALLAALAPARTDRGSDGGGDRGGGRVALTAPGA 420  
Db 347 ASCVAVIEHSHWAHRAFPAPFLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395

Qy 421 ADAPGAGATAGIHWSQLLYOIGTWLDDSEALHPLGNVAKSS 462  
Db 396 TEARGAPTAGIHWSQLLYHIGTWLDDSETMHPHGNVAKSS 437

RESULT 40  
US-09-969-520A-9  
; Sequence 9, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Philip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04

; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-9

Query Match 83.2%; Score 2054; DB 9; Length 437;  
Best Local Similarity 87.0%; Pred. No. 2.9e-171;  
Matches 402; Conservative 8; Mismatches 26; Indels 26; Gaps 3;

Qy 1 MILLARCLLLVSSILVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
Db 2 LLLARCFVLVILASSLLVCFGLACGPGRGKRRHPKLTALAYAQFIPNVAEKTIGASG 61

Qy 61 RYEGKISRNSRPFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQPGV 120  
Db 62 RYEGKITNSRPFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISSVMNQPGV 121

Qy 121 KLRTGWDGDDGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWYYESKAH 180  
Db 122 KLRTGWDGDDGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWYYESKAH 181

Qy 181 IHCSVKAENSVAAGSGCGFFGSAIVHLEQGTGLVKDLSGDRVLAADQGRLLYSDFLT 240  
Db 182 IHCSVKAENSVAAGSGCGFFGSAIVHLEQGTGLVKDLSGDRVLAADQGRLLYSDFLT 241

Qy 241 FLDRDGAKKVYVIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300  
Db 242 FLDRDGAKKVYVIETLEPRERLLTAAHLLFVAPHND-----SGPTPG 286

Qy 301 PRALFASVRPGQVYVVAERDGRLLPAAVHVSVTLSEEAAGAYAPLTAQGTILINRVL 360  
Db 287 PSALFASVRPGQVYVVAERDGRLLPAAVHVSVTLSEEAAGAYAPLTAHGTILINRVL 346

Qy 361 ASCVAVIEHSHWAHRAFPAPFLAHALLAALAPARTDRGSDGGGDRGGGRVALTAPGA 420  
Db 347 ASCVAVIEHSHWAHRAFPAPFLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395

Qy 421 ADAPGAGATAGIHWSQLLYOIGTWLDDSEALHPLGNVAKSS 462  
Db 396 TEARGAPTAGIHWSQLLYHIGTWLDDSETMHPHGNVAKSS 437

RESULT 41  
US-09-969-520A-4  
; Sequence 4, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Philip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-4

Query Match 83.1%; Score 2052; DB 9; Length 437;



QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDSGGDRGGGGEVALTAPGA 420  
 Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
 QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSS 462  
 Db 396 TEARGABTAGIHWSQLLYHIGTWLLDSETMPLGMVAKSS 437

## RESULT 44

US-09-969-520A-7  
 ; Sequence 7, Application US/09969520A  
 ; Patent No. US20020177163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
 ; APPLICANT: BEACHY, Philip A.  
 ; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
 ; FILE REFERENCE: JHU1670-1  
 ; CURRENT APPLICATION NUMBER: US/09/969,520A  
 ; CURRENT FILING DATE: 2002-06-04  
 ; PRIOR APPLICATION NUMBER: US 60/235,153  
 ; PRIOR FILING DATE: 2000-09-22  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Altered sonic hedgehog protein  
 US-09-969-520A-7

Query Match 82.5%; Score 2036; DB 9; Length 437;  
 Best Local Similarity 86.1%; Pred. No. 1.1e-169;  
 Matches 398; Conservative 10; Mismatches 28; Indels 26; Gaps 3;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFINVAEKTILGASG 60  
 Db 2 LLLARCLFLVILASSLLVCSGLACGPGRGFGKRRHPKLTALAYAFINVAEKTILGASG 61  
 QY 61 RYEGKISRNSERFKELTPNYPNPIIFKDEBNTGADRLMTQRCXKLNALAI SVMNQPGV 120  
 Db 62 RYEGKITRNSERFKELTPNYPNPIIFKDEBNTGADRLMTQRCXKLNALAI SVMNQPGV 121  
 QY 121 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWYYESKAH 180  
 Db 122 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRAKYGMRLARLAVEAGFDWYYEAAAH 181  
 QY 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 182 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLRPGDRVLAADDQGRLLYSDFLT 241  
 QY 241 FLDRDGAKKVYVIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPFGSGALG 300  
 Db 242 FLDRDEGAKKVYVIETLEPRERLLTAAHLLFVAPHND-----SGPTFG 286  
 QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
 Db 287 PSALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSREEAAGAYAPLTAHGTILINRVL 346  
 QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDSGGDRGGGGRVALTAPGA 420  
 Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
 QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSS 462  
 Db 396 TEARGABTAGIHWSQLLYHIGTWLLDSETMPLGMVAKSS 437

## RESULT 45

US-09-969-520A-5  
 ; Sequence 5, Application US/09969520A

; Patent No. US20020177163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
 ; APPLICANT: BEACHY, Philip A.  
 ; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
 ; FILE REFERENCE: JHU1670-1  
 ; CURRENT APPLICATION NUMBER: US/09/969,520A  
 ; CURRENT FILING DATE: 2002-06-04  
 ; PRIOR APPLICATION NUMBER: US 60/235,153  
 ; PRIOR FILING DATE: 2000-09-22  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Altered sonic hedgehog protein  
 US-09-969-520A-5

Query Match 81.8%; Score 2019; DB 9; Length 437;  
 Best Local Similarity 85.9%; Pred. No. 3.4e-168;  
 Matches 397; Conservative 8; Mismatches 31; Indels 26; Gaps 3;

QY 1 MLLARCLLLVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFINVAEKTILGASG 60  
 Db 2 LLLARCLFLVILASSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFINVAEKTILGASG 61  
 QY 61 RYEGKISRNSERFKELTPNYPNPIIFKDEBNTGADRLMTQRCXKLNALAI SVMNQPGV 120  
 Db 62 RYEGKITRNSERFKELTPNYPNPIIFKDEBNTGADRLMTQRCXKLNALAI SVMNQPGV 121  
 QY 121 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWYYESKAH 180  
 Db 122 KLRVTEGWDEGHHSEESLHYEGRAVDITTSDDRSKYGMRLARLAVEAGFDWYYESKAH 181  
 QY 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLT 240  
 Db 182 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLRPGDRVLAADDQGRLLYSDFLT 241  
 QY 241 FLDRDGAKKVYVIETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSGPFGSGALG 300  
 Db 242 FLDRDEGAKKVYVIETLEPRERLLTAAHLLFVAPHND-----SGPTFG 286  
 QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSSEAAGAYAPLTAQGTILINRVL 360  
 Db 287 PSALFASRVPRGQRYVVAERDGRLLPAVHVSVTLSREEAAGAYAPLTAHGTILINRVL 346  
 QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGSDSGGDRGGGGRVALTAPGA 420  
 Db 347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAQSA 395  
 QY 421 ADAPGAGATAGIHWSQLLYQIGTWLLDSEALHPLGMVAKSS 462  
 Db 396 TEARGABTAGIHWSQLLYHIGTWLLDSETMPLGMVAKSS 437

## RESULT 46

US-08-900-220C-10  
 ; Sequence 10, Application US/08900220C  
 ; Publication No. US20020045208A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Monica  
 ; APPLICANT: Mahanthappa, Nagesh K.  
 ; APPLICANT: Pang, Kevin  
 ; APPLICANT: Jin, Ping  
 ; TITLE OF INVENTION: Method of Treating Dopaminergic and  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: ONE POST OFFICE SQUARE



QY 359 VLASCVAVIEHNSWAHRAFPAPFRLAHALLAALAPARTDRGGSGGDRGGGRVALTAP 418  
 DB 348 VLASCVAVIEHNSWAHRAFPAPFRLAQGLAAL-----CP 381  
 QY 419 GAADAPGAGATAGIHWISOLLYOIGTWLSDSEALHPLGMVAKSS 462  
 DB 382 DGAIPTAATTTGIIHWISRLLYRIGSWLDGDLHPLGMVAPAS 425

RESULT 48  
 US-08-462-386D-8  
 ; Sequence 8, Application US/08462386D  
 ; Publication No. US20030186357A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ingham, Phillip W.  
 ; APPLICANT: McMahon, Andrew P.  
 ; APPLICANT: Tabin, Clifford J.  
 ; TITLE OF INVENTION: Vertebrate Tissue Pattern-Inducing  
 ; TITLE OF INVENTION: Proteins and Uses Related Thereto  
 ; NUMBER OF SEQUENCES: 47  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 State Street, Suite 510  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII(text)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/462,386D  
 ; FILING DATE: 5-JUNE-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/435,093  
 ; FILING DATE: 4-MAY-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/356,060  
 ; FILING DATE: 14-DEC-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/176,427  
 ; FILING DATE: 30-DEC-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vincent, Matthew P.  
 ; REGISTRATION NUMBER: 36,709  
 ; REFERENCE/DOCKET NUMBER: HM1-006CP3  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 425 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-462-386D-8

Query Match 73.2%; Score 1807; DB 8; Length 425;  
 Best Local Similarity 77.8%; Pred. No. 1.2e-149;  
 Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;  
 QY 1 MLLARCLLVVSSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 DB 4 MLLTRILLVGFICALLVSSGLTCGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 63  
 QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQWPGV 120  
 DB 64 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQWPGV 123  
 QY 121 KLRTVEGDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYESKAH 180

DB 124 KLRTVEGDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYESKAH 183  
 QY 181 IHCSVKAENSVAAKSGCGCPGSGATVHLBOGGTKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
 DB 184 IHCSVKAENSVAAKSGCGCPGSGATVHLBOGGTKLVKDLSPGDRVLAADDGRLLYSDFLT 243  
 QY 241 FLDRDDGAKKVFYVETREPRERLLTLTAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA 298  
 DB 244 FLDRDSSRKL FVETRETQPRARLLTLTAHLLFVAPQHNSGATG-----STSG----- 292  
 QY 299 LGPRALFASRVPRGQVYVVAERDGRRLLPAAVHSVTLSSEAGAYAPLTAOGTILNR 358  
 DB 293 ---QALFASNVKPGQVYVVLGE--GGQQLLPASVHSVLSREESAGAYAPLTAOGTILNR 347  
 QY 359 VLASCVAVIEHNSWAHRAFPAPFRLAHALLAALAPARTDRGGSGGDRGGGRVALTAP 418  
 DB 348 VLASCVAVIEHNSWAHRAFPAPFRLAQGLAAL-----CP 381  
 QY 419 GAADAPGAGATAGIHWISOLLYOIGTWLSDSEALHPLGMVAKSS 462  
 DB 382 DGAIPTAATTTGIIHWISRLLYRIGSWLDGDLHPLGMVAPAS 425

RESULT 49  
 US-09-021-660A-34  
 ; Sequence 34, Application US/09021660A  
 ; Patent No. US20010041668A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baron, M.  
 ; APPLICANT: Farrington, S.  
 ; APPLICANT: Belausoff, M.  
 ; TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR  
 ; TITLE OF INVENTION: GROWTH  
 ; FILE REFERENCE: HU/P-PO1-060  
 ; CURRENT APPLICATION NUMBER: US/09/021,660A  
 ; CURRENT FILING DATE: 2001-08-27  
 ; PRIOR APPLICATION NUMBER: 60/037,513  
 ; PRIOR FILING DATE: 1997-02-10  
 ; PRIOR APPLICATION NUMBER: 60/049,763  
 ; PRIOR FILING DATE: 1997-06-16  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 34  
 ; LENGTH: 425  
 ; TYPE: PRT  
 ; ORGANISM: Gallus gallus  
 ; US-09-021-660A-34

Query Match 73.2%; Score 1807; DB 9; Length 425;  
 Best Local Similarity 77.8%; Pred. No. 1.2e-149;  
 Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;  
 QY 1 MLLARCLLVVSSLLVCSGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 60  
 DB 4 MLLTRILLVGFICALLVSSGLTCGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGASG 63  
 QY 61 RYEGKISRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQWPGV 120  
 DB 64 RYEGKITRNSERFKELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQWPGV 123  
 QY 121 KLRTVEGDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYESKAH 180  
 DB 124 KLRTVEGDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYESKAH 183  
 QY 181 IHCSVKAENSVAAKSGCGCPGSGATVHLBOGGTKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
 DB 184 IHCSVKAENSVAAKSGCGCPGSGATVHLBOGGTKLVKDLSPGDRVLAADDGRLLYSDFLT 243  
 QY 241 FLDRDDGAKKVFYVETREPRERLLTLTAHLLFVAP-HNDS-ATGEPEASSGSGPPSGGA 298  
 DB 244 FLDRDSSRKL FVETRETQPRARLLTLTAHLLFVAPQHNSGATG-----STSG----- 292



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 18:28:04 ; Search time 20 Seconds  
(without alignments)  
2284.549 Million cell updates/sec

Title: US-09-883-848A-15  
Perfect score: 2469  
Sequence: 1 MLLARCLLVVSSLLVCS.....GMAVKSSXRGGAGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2062	83.5	437	2 A49425	Sonic hedgehog pro
2	2060	83.4	437	2 B53193	hedgehog homolog v
3	1807	73.2	425	2 A49424	patterning protein
4	1506	61.0	418	2 A53193	hedgehog homolog v
5	1498	60.7	444	2 S56765	morphogen Xhh prec
6	1459	59.1	415	2 A49426	sonic hedgehog gen
7	1183	47.9	396	2 B49425	Desert hedgehog pr
8	1027.5	41.6	336	2 C49425	Indian hedgehog pr
9	890	36.0	471	2 A46400	segment polarity p
10	449	18.2	94	2 G02735	desert hedgehog -
11	159.5	6.5	615	2 T29550	hypothetical prote
12	149	5.0	1207	2 T23754	hypothetical prote
13	138.5	5.6	1226	2 T44045	hypothetical prote
14	135.5	5.5	484	2 T4504	hypothetical prote
15	134.5	5.4	481	2 T27665	hypothetical prote
16	129	5.2	1147	1 MWXIB	myosin heavy chain
17	128	5.2	629	2 T19563	hypothetical prote
18	126.5	5.1	1021	2 T23522	hypothetical prote
19	123.5	5.0	2639	2 T31328	fibroin - Chinese
20	121	4.9	868	2 T32281	hypothetical prote
21	119	4.8	318	2 T35213	probable secreted
22	118.5	4.8	702	2 G01840	T-box protein 2 -
23	116	4.7	1329	2 E70917	hypothetical glyci
24	115.5	4.7	940	2 H87250	(protein-pil) urid
25	115	4.7	486	2 H87311	hypothetical prote
26	114	4.6	741	2 G70917	hypothetical glyci
27	111	4.5	333	2 A35065	homeotic protein E
28	111	4.5	1298	1 EDBE75	immediate-early pr
29	110	4.5	462	2 T35852	probable UDP-N-ace

30	107.5	4.4	701	1 S46458	transcription fact
31	107	4.3	775	1 EDBE11	immediate-early pr
32	106.5	4.3	482	2 E70554	hypothetical prote
33	106	4.3	482	2 JC7583	basic helix-loop-h
34	105.5	4.3	490	2 T09084	phosphatidylinosit
35	105.5	4.3	841	2 C87331	ISC22, transposase
36	105.5	4.3	1324	2 T14070	peptide synthetase
37	105.5	4.3	7463	2 T36248	hypothetical prote
38	105	4.3	460	2 T33110	glycerol-3-phospha
39	104.5	4.2	510	2 A55207	monomeric polyketid
40	104	4.2	402	1 S25077	acidic nuclear pro
41	104	4.2	1063	2 A40253	protein HMWPI - Ye
42	104	4.2	3161	2 T30342	homeotic protein C
43	103.5	4.2	975	2 S33121	probable PE protei
44	102.5	4.2	413	2 E70661	hypothetical prote
45	102.5	4.2	479	2 F70573	hypothetical prote
46	102.5	4.2	481	2 T27975	hypothetical prote
47	102	4.1	1145	2 B75625	hypothetical prote
48	102	4.1	1209	2 T00373	hypothetical prote
49	101.5	4.1	506	2 B43339	hypothetical prote
50	101.5	4.1	664	2 C70584	probable serine-th
51	101.5	4.1	777	2 C86454	hypothetical prote
52	101.5	4.1	1958	2 B40505	hypothetical prote
53	100.5	4.1	313	2 T15855	hypothetical prote
54	100.5	4.1	430	2 JC7379	levansucrase (EC 2
55	100.5	4.1	2647	2 A37098	gelation factor AB
56	100	4.1	249	2 S41374	single chain Fv an
57	100	4.1	825	1 EDBE11	immediate-early pr
58	99.5	4.0	361	2 G70682	hypothetical glyci
59	99.5	4.0	598	2 T42070	protein serine/thr
60	99.5	4.0	874	2 B70945	hypothetical prote
61	99	4.0	244	2 T36357	probable membrane
62	99	4.0	758	2 A83121	probable two-compo
63	99	4.0	831	2 H84368	MCM / cell divisio
64	99	4.0	1022	2 T17406	developmental prot
65	98.5	4.0	867	2 S57795	probable deoxyribo
66	98	4.0	1061	2 A40609	OmpA-related prote
67	98	4.0	1376	1 VCBE06	major capsid prote
68	98	4.0	2129	2 T14182	fixC protein - Myc
69	97.5	3.9	731	2 C70974	hypothetical glyci
70	97	3.9	519	2 A54590	GAGA transcription
71	97	3.9	808	2 F84038	phenylalanyl-tRNA
72	97	3.9	1010	2 I40329	brkA ptotein - Bor
73	97	3.9	1096	2 C87263	hypothetical prote
74	97	3.9	1541	2 T02831	AAA protein J4171
75	96.5	3.9	341	2 C83578	hypothetical prote
76	96.5	3.9	410	2 JC7584	basic helix-loop-h
77	96.5	3.9	443	2 D87287	pmbA protein limpo
78	96.5	3.9	495	1 S31223	transcription fact
79	96.5	3.9	604	2 S25203	srnR protein - Str
80	96.5	3.9	930	2 T33180	hypothetical prote
81	96.5	3.9	1487	1 EDBE11	immediate-early pr
82	96.5	3.9	1487	1 EDBE11	155K transcription
83	96	3.9	839	2 F75518	hypothetical prote
84	96	3.9	1017	2 F82500	vgrG protein VCA01
85	96	3.9	1137	2 JC5950	integrin alpha-7 c
86	96	3.9	3163	2 A70514	probable polyketid
87	96	3.9	3163	2 A70514	versinibactin bio
88	95.5	3.9	431	2 S09824	hypothetical glyci
89	95.5	3.9	461	2 T51044	hypothetical prote
90	95.5	3.9	475	2 A43915	related to spore c
91	95.5	3.9	497	2 F82985	homeotic protein e
92	95.5	3.9	756	2 T04187	conserved hypothet
93	95.5	3.9	908	2 A31280	subtilisin-like pr
94	95.5	3.9	908	2 A31280	sarcosylamin precu
95	95	3.8	419	2 AP2274	tyrosyl-tRNA synth
96	95	3.8	529	2 H84049	hypothetical prote
97	95	3.8	558	2 D69067	succinate dehydrog
98	95	3.8	577	2 AH2908	hypothetical prote
99	95	3.8	588	2 G97683	prsd protein (U891
100	95	3.8	767	2 A35645	major surface prot
101	95	3.8	825	2 JC4163	DNA-binding protei
102	95	3.8	894	2 C86756	prophage p12 prote



```

241 FLDRDDGAKVYVYIETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
242 FLDRDDGAKVYVYIETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 286
301 PRALFASVRPQGVVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 360
287 PSLFASVRPQGVVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 346
361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDGGSDGGGGRGGRVALTAP-G 419
347 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDGGSDGGGGRGGRVALTAP-G 394
420 AADAPGAGATAGIHWSQLLYQIGTWTLLDSEALHPLGMVAKSS 462
395 VAERAGGPPAGIHWSQLLYHIGTWTLLDSEALHPLGMVAKSS 437

RESULT 3
A49424
patterning protein sonic hedgehog precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
R:Riddle, R.D.; Johnson, R.L.; Laufer, E.; Tabin, C.
Cell 75, 1401-1416, 1993
A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.
A:Reference number: A49424; MUID:94094333; PMID:8269518
A:Accession: A49424
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-425 <RID>
A:Cross-references: GB:L28099; NID:G453526; PIDN:AAA72428.1; PID:G453527
C:Superfamily: sonic hedgehog protein
F:1-26/Domain: signal sequence #status predicted <SIG>

Query Match 73.2%; Score 1807; DB 2; Length 425;
Best Local Similarity 77.8%; Pred. No. 4.8e-128; Indels 44; Gaps 6;
Matches 361; Conservative 22; Mismatches 37;

QY 1 MLLARCLLLVLSLVCSGLACGPGRGKRRHPKLTPLAYKQIFPNVAEKTILGASG 60
Db 4 MLLTRILLVFCALLVSSGLTCGPGRGKRRHPKLTPLAYKQIFPNVAEKTILGASG 63
QY 61 RYEGKISNSRPFKELTPNYPNDIIFKDEENTGADRLMTORCKDKNALAISVNNOWPGV 120
Db 64 RYEGKITRNSRPFKELTPNYPNDIIFKDEENTGADRLMTORCKDKNALAISVNNOWPGV 123
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180
Db 124 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 183
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGDRVLAADDGRLLYSDFLT 240
Db 184 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGDRVLAADDGRLLYSDFLT 243
QY 241 FLDRDDGAKVYVYIETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGA 298
Db 244 FLDRDDGAKVYVYIETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGA 292
QY 299 LGPRALFASVRPQGVVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAQGTILINR 358
Db 293 ---QALFASNVKPGQVVYVLGE--GGQQLLPASVHVSLSREAGAYAPLTAQGTILINR 347
QY 359 VLACVAVIEHSHWAHRAFAFRLAHALLAALAPARTDGGSDGGGGRGGRVALTAP 418
Db 348 VLACVAVIEHSHWAHRAFAFRLAHALLAALAPARTDGGSDGGGGRGGRVALTAP 381
QY 419 GAADAPGAGATAGIHWSQLLYQIGTWTLLDSEALHPLGMVAKSS 462
Db 382 DGAIPATAATTTGIHWYRSLLYRIGSVWLDGDLHPLGMVAPAS 425
```

## RESULT 4

```

A53193
hedgehog homolog vhh-1 - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
C:Accession: A53193
R:Roelink, H.; Augburger, A.; Heemkerk, J.; Kozh, V.; Norlin, S.; Ruiz i Altaba, A.;
Cell 76, 761-775, 1994
A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog
A:Reference number: A53193; MUID:94170375; PMID:8124714
A:Accession: A53193
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <ROE>
A:Cross-references: GB:L27585; NID:G452159; PID:G452160
C:Superfamily: sonic hedgehog protein
```

Query Match 61.0%; Score 1506; DB 2; Length 418;

Best Local Similarity 64.5%; Pred. No. 1.9e-105; Indels 50; Gaps 5;  
Matches 300; Conservative 47; Mismatches 68;

```

QY 1 MLLARCLLLVLSLVCSGLACGPGRGKRRHPKLTPLAYKQIFPNVAEKTILGASG 60
Db 1 MLLTRILLVLSLVCSGLACGPGRGKRRHPKLTPLAYKQIFPNVAEKTILGASG 60
QY 61 RYEGKISNSRPFKELTPNYPNDIIFKDEENTGADRLMTORCKDKNALAISVNNOWPGV 120
Db 61 RYEGKITRNSRPFKELTPNYPNDIIFKDEENTGADRLMTORCKDKNALAISVNNOWPGV 120
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180
Db 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH 180
QY 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGDRVLAADDGRLLYSDFLT 240
Db 181 IHCSVKAENSVAAGSGCGFPGSATVHLEQGTGKLVKDLSPGDRVLAADDGRLLYSDFLT 240
QY 241 FLDRDDGAKVYVYIETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
Db 241 FLDRDDGAKVYVYIETREPRERLLITAAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 286
QY 301 PRALFASVRPQGVVVAERDGRRLPAAVHVSVTLSSEAGAYAPLTAQGTILINRVL 360
Db 287 ---AAYASSVRAGQKVMVY----DSSGQLKSVIVQRIYEEQSGFAPYTAHGTIVVDKIL 340
QY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDGGSDGGGGRGGRVALTAPGA 420
Db 341 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDGGSDGGGGRGGRVALTAPGA 375
QY 421 ADAPGAGAT---AGIHWSQLLYQIGTWTLLDSEALHPLGMVAKSS 462
Db 376 ---SSRSNATLQGEVHWYRSLLYQMGTLWLLDSNMLHPLGMSVNSS 418
```

## RESULT 5

```

S56765
morphogen Xhh precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S56765
R:Stolow, M.A.; Shi, Y.B.
Nucleic Acids Res. 23, 2555-2562, 1995
A:Title: Xenopus sonic hedgehog as a potential morphogen during embryogenesis and thymo
A:Reference number: S56765; MUID:95357169; PMID:7630736
A:Accession: S56765
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-444 <STO>
A:Cross-references: EMBL:L39213; NID:G790937; PIDN:AAC42227.1; PID:G790938
C:Superfamily: sonic hedgehog protein
```

Query Match 60.7%; Score 1498; DB 2; Length 444;  
Best Local Similarity 64.1%; Pred. No. 8.2e-105;

Qy	241	FLRDPDGAKKVYFVETREPREELLITAAHLLFVA	PHNDSATCEPEASSGSGPPSGCALG	300
Db	241	FLRDPDGAKKVYFVETREPREELLITAAHLLFVL	-----DNSTEDLHWT-----	286
Qy	301	PRALFASRVPGQGVVYVAERDGRLLLPANVISVTL	SEBACAYAPLTAQCTILINRVL	360
Db	287	--AAYASSVRAGQKVMW----	DDSGQLKSVIVQRIYTEQRGSFAPVTAHGTIVDRIIL	340
Qy	361	ASCYAVIEHSWAHRFAFPFLAHALLAALAPART	DGSDGGSGDGGGRRVALTAPGA	420
Db	341	ASCYAVIEDQGLAHAFAPARLYYYVSSFLSP	-KTPAVGPMRLYNRRGSGT-----TPGS	394
Qy	421	ADAPGAGATAGIHWYQLLYQIGTLLDSEALHPLGM	AV 459	
Db	395	C-----	HQMGTWLLDSNMLHPLGMSV	415

RESULT 7

B49425

Desert hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999

C:Accession: B49425

R:Echelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J. Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules. A:Reference number: A49425; MUID:94094334; PMID:7916661

A:Accession: B49425

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-396 <ECH>

A:Cross-references: GB:X76292; NID:g443941; PIDN:CAAS3924.1; PID:g443942

:Genetics:

Query Match	47.9%;	Score 1183;	DB 2;	Length 396;
Best Local Similarity	54.5%;	Pred. No. 3.3e-8;		
Matches 244;	Conservative	55;	Mismatches 87;	Indels 52; Gaps 10;
QY	3	LLARCLLLVLVSSLLVSGLAGCGPRG- FGKRRHPKK- LTPLAYKQFIPNVAETLGASG	60	
DB	7	LLPLCCL- ----ALLALSAQCGPGRGVGRRYVRKQLVPLLYKQFVPSMPERTLGASG	61	
QY	61	RYEKKISRNSEKPELPTNYPNDIIFKDEENTGADRLMTORCKDKALALATSVNQMPGV	120	
DB	62	PAEGRVTKGSRFRDLVPNYPNDIIFKDEENSGADRLMTERCKGRVNALALAVNMWPGV	121	
QY	121	KLRVTEGWDEGDHGHSESLHYEGRAVDITTSDDRKYGMRLARLAVEAGFDWVYSEKAH	180	
DB	122	KLRVTEGWDEGDHGHQAQDSLHYEGRALDITTSDRDNKYGLLARLAVEAGFDWVYSESNH	181	
QY	181	IHCYSKAEKNSVAASKGCGFPGSATVHLIEQGGTKLVKOLSPGDRVLADDQGGELLYSDFLT	240	
DB	182	IHSVYKADNLSAVRAGCGCFGNATVRLSRGEBKGRLLHREGDWVLAAADAGRVPTPVLL	241	
QY	241	FLDRDDGAKKYFVYVETREPRERLLTAAHLFLVAPHNDSATGPEASGSGP- PSGGAL	299	
DB	242	FLDRDLQRRASFVAVETERPRLKLLTPWHLVFAA- -----RGPAPAPGDF	286	
QY	300	GPRLFASRVPRPGORVYVVAERDGDRLLPAAVHSVTLSEAGNAYAPLTAGQTILINRV	359	
DB	287	AP- -VFARLRAGDSVLA- ---PGDALQPARVARVA- REEAVGVFAPLTAHGTLLVNDV	339	
QY	360	LASCVAVTEESWARRAFPRLAHALAALAPARTDRGSDGGDRGGGGRVALTAPG	419	
DB	340	LASCVAVLESQWARRAFAPRLLLHALCALL- -----PG	372	
QY	420	AADAPGAGATGIHWYSQLLIQIGTWLL	447	
DB	373	GAVQP- ----TCGMHYSRLLYELAEELM	395	

Db	182	IHVSVKADNSLAVRAGGCFPNQAIIVRLSGERKGLRELHRCGDWVLAADAAGRVPVTPVLL	244
QY	241	FLDRDDGAKKYFYVIETREPRERLLLTAHULLFVAPHNDSATGEPEASSGGSP--PSGAL	299
Db	242	FLDRDLQRRASFVAVETERPRKLLLTTPWHLVFAA-----RGPAPAGDF	286
QY	300	GPRLFASRVTPGQORVYVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAGGTILNRY	359
Db	287	AP--VFARLRAGOSVLA-----PGDALQPARVARVA--REEAVGVFAELTAHGILLVNDV	339
QY	360	LASCYAVITEESHWAFRAPFLRLAHALLAALAPARTDRGDSGGDRGGGGRVALTPG	419
Db	340	LASCYAVILESQWAFRAFAPLRLHALCALL-----PG	372
QY	420	AADAPGAGATAGIHWYSQLLYQIGTWLL	447
Db	373	GAVQP-----TGMHWYSRLLYELAEELM	395

## RESULT 8

C49425  
Indian hedgehog protein - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000  
C;Accession: C49425  
R;Chelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMahon, J.; 1417-1430, 1993  
A;Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the development of the mouse embryo.  
A;Reference number: A49425; MUID:94094334; PMID:7916661  
A;Accession: C49425  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-336 <ECH>  
A;Cross-references: GB:X76291  
C;Superfamily: sonic hedgehog protein

Query Match 41.6%; Score 1027.5; DB 2; Length 336;  
Best Local Similarity 54.6%; Pred. No. 1.3e-69;  
Matches 218; Conservative 39; Mismatches 77; Indels 65; Gaps 8;  
  
Qy 71 ERPKELTNNYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVNQNPGVKLVRTSGWDE 130  
Db 1 ERPKELTNNYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVNQNPGVKLVRTSGWDE 60  
  
Qy 131 DGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAHIHCSVKAENS 190  
Db 61 DGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVESKAHIHCSVKAENS 120  
  
Qy 191 VAAKSGGCGFPGSATVHLBOGGTKLVKDLSPGDRVLAADDQGRLLYSDFLTFLDRDDGAKK 250  
Db 121 AAATGGCGFPAQVRLENGERVALSAVKPGDRVLANGEDGTPTFSVLFLDRENRLR 180  
  
Qy 251 VEVYIETREPRERLLTAHLLFVAPHNDSATGPEASGSGPPSGGALGPALFASRVR 310  
Db 181 AFQVETQDPRRLALTPAHLFTA---DNHT-EPAAHF-----RATFASHVQ 224  
  
Qy 311 PGQRYVVAERDGRRLLPAAVHVSVTLSSEAAAGAYAPLTAQGTILINRVLASCVATEEH 370  
Db 225 PGQYVLV----SGVPGQLQPARVAAS-THVALGSYAPLTHRGTLVVEDVVAFCFAAVADH 279  
  
Qy 371 SWAHRAFAPEFLAHLA--AALAPARTDRGGSGGGGGRVALTAPGAADAPGAGAT 429  
Db 280 HLAQLAFWPLRFLPSLANGSWTPE-----304  
  
Qy 430 AGIHWYSQLLYQIGTWLDSSEALHPLGMVKSXSRRGAG 468  
Db 305 -GVHSYPMQLYRLGRLLEESTFHPGMS-----GAG 335

## RESULT 9

A46400  
segment polarity protein hedgehog - fruit fly (Drosophila melanogaster)  
N;Alternate names: hh protein  
C;Species: Drosophila melanogaster  
C;Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
C;Accession: A46400; JN0501; A43480  
R;Tabata, T.; Katon, S.; Kornberg, T.B.  
Genes Dev. 6, 2635-2645, 1992  
A;Title: The Drosophila hedgehog gene is expressed specifically in posterior compartment  
A;Reference number: A46400; MUID:94040725; PMID:1340474  
A;Accession: A46400  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-471 <TAB>  
A;Cross-references: GB:S66384; NID:9435848; PID:9435849  
A;Experimental source: Oregon-R  
A;Note: sequence extracted from NCBI backbone (NCBIN:138996, NCBI:P:138997)  
R;Tashiro, S.; Michie, T.; Higashijima, S.; Zengo, S.; Ishimaru, S.; Takahashi, F.; Orii, Y.; 183-189, 1993  
A;Title: Structure and expression of hedgehog, a Drosophila segment-polarity gene required for normal development of the wing.  
A;Reference number: JN0501; MUID:93185922; PMID:816882

A;Accession: JN0501  
A;Molecule type: mRNA  
A;Residues: 1-471 <TAS>  
A;Cross-references: GB:L05404  
A;Note: it is uncertain whether Met-1 or Met-51 is the initiator  
A;Note: intron positions were determined from partial DNA sequence  
R;Lee, J.J.; von Kessler, D.P.; Parks, S.; Beachy, P.A.  
Cell 71, 33-50, 1992  
A;Title: Secretion and localized transcription suggest a role in positional signaling f.  
A;Reference number: A43480; MUID:93008241; PMID:1394430  
A;Accession: A43480  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-471 <LEE>  
A;Cross-references: GB:L02793; NID:g157609; PID:g157610  
A;Note: sequence extracted from NCBI backbone (NCBI:P:115418)  
C;Comment: This protein is required for cell-cell communication.  
C;Genetics:

A;Gene: hh  
A;Cross-references: FlyBase:FBgn0004644  
A;Introns: 160/3; 248/1  
C;Superfamily: sonic hedgehog protein  
C;Keywords: transmembrane protein  
P;62-82/Domain: transmembrane #status predicted <TM>

Query Match 36.0%; Score 890; DB 2; Length 471;  
Best Local Similarity 44.3%; Pred. No. 4.2e-59;  
Matches 202; Conservative 65; Mismatches 137; Indels 52; Gaps 10;

Qy 6 RCL-----LLVLVSLLVCSGLACGFGKRRHPKLTPLAYKQFIPNVAEKTG 57  
Db 59 RCLRLTSLVALLLIVLPMVFSAPHS CGFGRLGHR-ARNLYPLVLKQTIPNLSEYTN 117  
  
Qy 58 ASGRVEGKISNSERPKELTNNYNDPIIFKDEENTGADRLMTORCKDKLNALAI SVNQN 117  
Db 118 ASGPLEGVIRDSRPFKOLVPNNYNDPIIFRDEEGTGADRLMSKCKEKNLVLSVNNW 177  
  
Qy 118 PGVKLVRTGEMDGHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYVES 177  
Db 178 PGIRLLVTESWEDVHGQESLHYEGRAVTATSDQSKYGMARLAVEAGFDWVYVS 237  
  
Qy 178 KAHHCSCVKAENSVAAGSGGCGFPGSATVHLBOGGTKLVKDLSPGDRVLAADDQGRLLYS 237  
Db 238 RSHYCSVKSSISSHVHGCTPPESTALLESVGKPLGELSIGDRVLSMTANGOAVSE 297  
  
Qy 238 FLTFELDRDGAKKVFTETREPRERLLTAHLLFVAPHNDSATGPEASGSGPPSGG 297  
Db 298 VILFMDRNLEQNFQVLT-DGGAVLTVPALYVS-----WQESQKLT----- 342  
  
Qy 298 ALGPALFASRVRPGQRYVVAERDGRRLLPAAVHVSVTLSSEAAAGAYAPLTAQGTILIN 357  
Db 343 -----FVFADRIEKNQVLVRDVTGELR--PQRYVVKVG-SVRSKGVVAPLTREGTIVN 394  
  
Qy 358 RVLASCVAVIEHSHWRAFAFRLAHALLAALAPARTDRGGSGGGGRVALT 417  
Db 395 SVAASCYAVINSQSLAHWGLAPMLLSTLEAWL-PAKEQ-----LHSS 436  
  
Qy 418 PGAAAPGACAGATAGIHWYSQLLYQIGTWLDSSEALH 453  
Db 437 PKWVSS--AQQNGIHWYANALYKVKDYVLPQSWRH 470

## RESULT 10

G02735  
desert hedgehog - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 11-Jan-2000  
C;Accession: G02735  
R;Drummond, I.A.  
submitted to the EMBL Data Library, June 1996  
A;Reference number: H01643  
A;Accession: G02735  
A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-94 <DRU>  
A:Cross-references: EMBL:U59748; NID:g1401271; PID:g1401272  
C:Genetics:  
A:Gene: hDRH  
C:Superfamily: sonic hedgehog protein

Query Match 18.2%; Score 449; DB 2; Length 94;  
Best Local Similarity 84.0%; Pred. No. 6.9e-27;  
Matches 79; Conservative 13; Mismatches 2; Indels 0; Gaps 0;

QY 84 IIFKDEBNTGADRLMTQRCXKDKLNALAIWMQWPGVKLRVTGEGWDEGHSEESLHYEG 143  
Db 1 IIFKDEBNTGADRLMTQRCXKDKLNALAIWMQWPGVKLRVTGEGWDEGHSEESLHYEG 60

QY 144 RAVDITTSRDRSKYGMARLAVRQGFDMVYYES 177  
Db 61 RALDITTSRDRSKYGMARLAVRQGFDMVYYES 94

RESULT 11  
T29550  
hypothetical protein ZK377.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
R:Nhan, M.; Hawkins, J.  
A:Accession: T29550  
Submitted to the EMBL Data Library, February 1997  
A:Description: The sequence of C. elegans cosmid ZK377.  
A:Reference number: Z20639  
A:Accession: T29550  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-615 <NHA>  
A:Cross-references: EMBL:U98183; PIDN:AAE52656.1; GSPDB:GN00028; CESP:ZK377.1  
A:Experimental source: strain Bristol N2; clone ZK377  
C:Genetics:  
A:Gene: CESP:ZK377.1  
A:Map position: X  
A:Introns: 46/1; 70/3; 157/2; 176/1; 245/2; 314/3; 402/3; 499/2

Query Match 6.5%; Score 159.5; DB 2; Length 615;  
Best Local Similarity 28.3%; Pred. No. 0.00044;  
Matches 52; Conservative 33; Mismatches 74; Indels 25; Gaps 7;

QY 198 CFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLFLDRDDGAKKVFYIET 257  
Db 416 CFPGDMVEVETD-GIKMIKDKIGDKVLSW-DEAFVTYSPVIMFLHKKRDEEIAEFNLLET 473

QY 258 REPRERLLTAAHLLVFAPHNDSATGPEASGSGPFGALGPRALFASRVPRGORVYV 317  
Db 474 ANGHS-IKLTNDHLIYV---SDCETRS-----DLKLVAAKEVKMDDCIHV 514

QY 318 VAERDGRRLLPAAVHVSITSEBAAGAYAPLTACGTILNRLVLAACVAVIEHSHWAHRAF 377  
Db 515 TTD---SNVVIKKVSKISKIVIE-TGIYSPLTSGDIIVNRLVLAACHSNLAKLSQOOTFF 570

QY 378 APPR 381  
Db 571 SLVK 574

RESULT 12  
T23754  
hypothetical protein T05C12.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
A:Accession: T23754; T24513  
R:Thomas, K.  
Submitted to the EMBL Data Library, June 1995  
A:Reference number: Z19793  
A:Accession: T23754  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-1207 <WIL>  
A:Cross-references: EMBL:Z49968; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10  
A:Experimental source: clone M10  
R:Burton, J.  
Submitted to the EMBL Data Library, October 1995  
A:Reference number: Z19901  
A:Accession: T24513  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1207 <WIL>  
A:Cross-references: EMBL:Z66500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10  
A:Experimental source: clone T05C12  
C:Genetics:  
A:Gene: CESP:T05C12.10  
A:Map position: 2  
A:Introns: 31/3; 87/2; 141/3; 180/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143/1;

Query Match 6.0%; Score 149; DB 2; Length 1207;  
Best Local Similarity 26.5%; Pred. No. 0.0065;  
Matches 56; Conservative 37; Mismatches 82; Indels 36; Gaps 9;

QY 184 SVKAENSVAAKSGG-----CFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDF 238  
Db 942 AVLADITFGACAGAGGSRNCFADSILV-TTGTGQKMDQLQIGDYVLVPSSGNVLKYEV 1000

QY 239 LTFDRDDGAKKVFYIETREPRERLLTAAHLLVFAPHNDSATGPEASGSGPFGSGGA 298  
Db 1001 EMFYHREPKRTNFVLYTKSGR-KLSLTGRHLLPVAEC-----QVEQYTNPNPDGID 1052

QY 299 LGPR-ALFASRVPRGORVYVVAERDGRRLLPAAVHVSITLSEBA-----AGAYAPLT 349  
Db 1053 VAMRESKYAEKARGECVLSIDE-SGB-----VIADEIVRGRVTNVIYSPMT 1100

QY 350 AQGTILNRLVLAACVAVIEHSHWAHRAF 380  
Db 1101 VEGSLIVDGVLSGCFSHLESHS-AHKLIIDF 1130

RESULT 13  
T24045  
hypothetical protein R08B4.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
A:Accession: T24045  
R:White, S.  
Submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19834  
A:Accession: T24045  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1226 <WIL>  
A:Cross-references: EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1  
A:Experimental source: clone R08B4  
C:Genetics:  
A:Gene: CESP:R08B4.1  
A:Map position: X  
A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 5.6%; Score 138.5; DB 2; Length 1226;  
Best Local Similarity 24.4%; Pred. No. 0.041;  
Matches 50; Conservative 38; Mismatches 68; Indels 49; Gaps 8;

QY 189 NSVAAKSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLLYSDFLFLDRDDGA 248  
Db 1012 SALVAATGACFSLDTWV-TTPTGKRMQDIDIGDYVLTADLE-KTYFTPTILTWIHFPEK 1069

QY 249 KKVFYIETREPRERLLTAAHLLF-----VAPHNDSATGPEASGSGGPS 295  
Db 1070 VQEPITMT-EYKTLRITSRHFMRNCKGSKYQYIKMLPHD-----GE----- 1113

QY 296 GGALGPRALFASRVPRGORVYVVAERDGRRLLPAAVHVSITLSEBAAGAYAPLTAAQGTIL 355  
A:Status: preliminary; translated from GB/EMBL/DBJ

Db 1114 -----AIFASDLEVGDCVVVLYKGYRQKLEITITRSV-----RTGIYSPLTNNGRII 1161

Qy 356 INRVLASCYAVIEHS-----NAH 374

Db 1162 VNDMLASCYSEIQNTLTQTTFFWAY 1186

RESULT 14

T34504

hypothetical protein ZK1290.12 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T34504

R:Taich, A.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid ZK1290.

A:Reference number: Z21535

A:Accession: T34504

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-484 <TAI>

A:Cross-references: EMBL:U21308; PIDN:AAB93321.1; GSPDB:GN00020; CESP:ZK1290.12

A:Experimental source: strain Bristol N2; clone ZK1290

C:Genetics:

A:Gene: CESP:ZK1290.12

A:Map position: 2

A:Introns: 66/3; 124/2; 181/2; 392/1

Query Match 5.5%; Score 135.5; DB 2; Length 484;

Best Local Similarity 24.5%; Pred. No. 0.021;

Matches 46; Conservative 42; Mismatches 65; Indels 35; Gaps 8;

Qy 191 VAAKSGCGFPGSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLFLDRDCAKK 250

Db 283 VASGVACFTGNSKV-MTPAGEKSWADLSVGMVMTY-EYKMTYTRVASWHLRLPDTKA 340

Qy 251 VFYVETPREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALGPRALFASRV 310

Db 341 AFIKL-TTEQAIIDMTPOHFIYA---NCVTEME-----LVYAE DMT 380

Qy 311 PGQRYVVAERDGRRLPAAVHVSITLSEEA-----AGAYAPLTAQGTILINRVLASCYAV 366

Db 381 IGDCLVW---KENELVW-----TTISEKSTFVETGVYAPMTETGDLIVDDVVASCHNV 431

Qy 367 IEHSWAH 374

Db 432 VKANTLSH 439

RESULT 15

T27665

hypothetical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27665

R:Basham, V.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z20401

A:Accession: T27665

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 <WIL>

A:Cross-references: EMBL:Z81142; PIDN:CAB03509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match 5.4%; Score 134.5; DB 2; Length 481;

Best Local Similarity 24.3%; Pred. No. 0.024;

Matches 45; Conservative 34; Mismatches 67; Indels 39; Gaps 6;

Qy 198 CPGSATVHLEQGGTKLVKOLSPGDRVLAADDQGRLLYSDFLFLD-----RDDGAKKV 252

Db 314 CFPNDVAVVYKAVKRMDELEIGDWVEALDENG-----EDITLPVKYWLHRDPEQAE 368

Qy 253 YVETPREPRERLLTAHLLFVAPHNDSATGPEASSGSGPPSGGALGPRALFASRV 312

Db 369 FLEFSLDNGETFTLTKELVY-----TTECRQNS-----SELKISWESISAGKNVAG 415

Qy 313 QRYVVAERDGRRLPAAVHVSITLSE-----EAQAYAPLTAQGTILINRVLASCYA 365

Db 416 DCFYLAQSE-----ALTKYRLVEILDIKRVKKTGIYAPMTSQGHLVKNKHTSCHS 466

Qy 366 VIEEH 370

Db 467 EVDHH 471

RESULT 16

MXAXIB

myosin heavy chain IB - Acanthamoeba castellanii

N:Contains: myosin ATPase (EC 3.6.4.1)

C:Species: Acanthamoeba castellanii

C>Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 19-Apr-2002

C:Accession: JQ0095; B34448

R:Jung, G.; Schmidt, C.J.; Hammer III, J.A.

Gene 82; 269-280, 1989

A:Title: Myosin I heavy-chain genes of Acanthamoeba castellanii: cloning of a second gene

A:Reference number: JQ0095; MUID:90060816; PMID:2511079

A:Accession: JQ0095

A:Molecule type: DNA

A:Residues: 1-1147 <JUN>

A:Cross-references: GB:M30780

A>Note: this organism expresses at least three isoforms of myosin I heavy-chain, encode of gene MIB, whereas the protein previously identified as MIB is the product of gene MI R:Brzeska, H.; Lynch, T.J.; Martin, B.; Korn, E.D.

J. Biol. Chem. 264, 19340-19348, 1989

A:Title: The localization and sequence of the phosphorylation sites of Acanthamoeba myo

A:Reference number: A34448; MUID:90037074; PMID:2530230

A:Accession: B34448

A:Molecule type: protein

A:Residues: 538-550, 'X', 552-559 <BR>

C:Comment: In this protein, the coiled-coil rod-like region found in many myosin heavy protein is globular and does not self-assemble into filaments.

C:Genetics:

A:Gene: MIB

A:Introns: 1/3; 39/3; 102/2; 135/3; 183/3; 212/1; 291/3; 379/3; 492/3; 617/2; 649/3; 78

C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homol

C:Keywords: actin binding; ATP; hydrolase; nucleotide binding; P-loop; phosphoprotein;

F:12-664/Domain: myosin motor domain homology <MMOT>

F:103-110/Region: nucleotide-binding motif A (P-loop)

F:552-573/Region: actin binding #status predicted

F:598-1147/Domain: carboxyl-terminal <CTD>

F:598-889/Region: basic

F:910-1094/Region: alanine/glycine/proline-rich

F:1097-1144/Domain: SH3 homology <SH3>

F:109/Binding site: ATP (Lys) #status predicted

F:315/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 5.2%; Score 129; DB 1; Length 1147;

Best Local Similarity 21.3%; Pred. No. 0.19;

Matches 104; Conservative 55; Mismatches 180; Indels 150; Gaps 21;

Qy 44 YKQFINVAEKTGLGASRGYEGKISNSERFKELTNNYNPDIIFKD-----EENTGADRLMT 99

Db 591 YLGLENNVRVRAGFA--YRAEFDFRLRYKKLSP-----KTWGIWGEWSGAPK--- 637

Qy 100 QRCKDKNALAISVMNQPGVKLVRTGWDGHHSESL-----HYEG 143

Db 638 DGCQTUNDLGLDT-SQWLGSKVFIRYPETLFLEELCDKDYDCTLRIOKANRWKS 696

Qy 144 R-----AVDITTSRDRSKYGMALAEVAFDQWVYKSKAHIIHCSVKAENSVAAK 194

Db 697 RXHOLEQRKMAADLLKGGKQRHSWRKY-----FDYINYDANYPLQDCVR--SSGRDK 750

RESULT 18

Db 2266 SGGYGGYGGSDSAAAAAAAAAASGAGGAGGYGGYG----GYGSYGSDSAAAAAAAAA 2321

```
QY 458 AVKSSXSRGAGGG 470
Db 2322 AAAGSGAGGVGGG 2334

RESULT 20
T22281
hypothetical protein F46B3.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T22281
R/Ainscouth, R.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19541
A/Accession: T22281
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-868 <WTL>
A/Cross-references: EMBL:Z81540; PIDN:CAB04405.1; GSPDB:GN00023; CESP:F46B3.5
A/Experimental source: clone F46B3
C/Genetics:
A/Gene: CESP:F46B3.5
A/Map position: 5
A/Introns: 45/3; 200/1; 235/3; 436/3; 561/3; 581/3; 696/3; 742/1; 826/3; 846/3

Query Match 4.9%; Score 121; DB 2; Length 868;
Best Local Similarity 26.7%; Pred. No. 0.54;
Matches 47; Conservative 32; Mismatches 73; Indels 24; Gaps 8;

QY 196 GCGPFGSATVHLEGGTKLVKDLSPGDRVLAADQDGLLYSDFLTFELDRDDGAKVYVI 255
Db 662 GCGF-SSDTLVTPSGKKRMDIDVDGYVLTARNVKTHTPTVTLWIHRESEKLEBFLTI 719

QY 256 ETRPRERLILTAHLLFVAPHNDATG---EPEASSGSGPPSGGALGPRALPASVRPG 312
Db 720 -TTERGTQLTPLHFVIRTKCNESSEFLKILPE-----NHEAILASVLEIG 765

QY 313 QRVVVAERDGRLLPAAVHVSILTSEEAAGAYAPLTAQQTILINRVLASCVAVIE 369
Db 766 DCV-ILTENTKFRQ---EKINOTTRGLK-TGIYSPLTKNGRIIIVNDMLASCYSEVQ 816

RESULT 21
T35213
probable secreted protein - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C/Accession: T35213
R/Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A/Reference number: Z21572
A/Accession: T35213
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-318 <SES>
A/Cross-references: EMBL:AL031515; PIDN:CAA20618.1; GSPDB:GN00070; SCODB:SC5C7.06
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCODB:SC5C7.06

Query Match 4.8%; Score 119; DB 2; Length 318;
Best Local Similarity 26.3%; Pred. No. 0.21;
Matches 59; Conservative 19; Mismatches 100; Indels 46; Gaps 9;

QY 276 PHNDATGEPASSGSGPPSGGALGPRALFASVRVPGQRTVVVAERDGRLLPAAVHSV 335
Db 46 PADDSAAVPPSAEAEAPPSSAQAAAR-VSPSDAGPGDTVTV-----SV 88

QY 336 TLSEEAAGAYAPLTA-----QGTILINRVLASCVAVIEHWSWAHRAFPAPRLAHALLAA 389
Db 89 SCGPTGGGAPASLDATSAFAEGTVALRKVADDACTA---SGPAYRGTAATAAEDFAAE 145

QY 390 LAPARTDRGGDSGGGGRGGRVALTAPGAADAPAGAGAT-----AGIHWYSQL--LYQI 442
Db 146 RPPADITDPG--GATDEGANPEGATPGSATDPGCVATDGGPSAGTEGMAETGDSAEA 203

QY 443 G-TWLLDSEALHPLG-----MAVKSSXSRGAGGAREGA 475
Db 204 GDDWTVDGACPDASGGGPGPWSATNSVPESGAGAGAGAAQPA 247

RESULT 22
G01840
T-box protein 2 - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 23-Jul-1999
C/Accession: G01840
R/Campbell, C.E.
submitted to the EMBL Data Library, May 1995
A/Reference number: G08602
A/Accession: G01840
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-702 <CAM>
A/Cross-references: EMBL:U28049; NID:G924927; PIDN:AAA73861.1; PID:G924928
C/Genetics:
A/Gene: GDB:TBX2
A/Cross-references: GDB:568496; OMIM:600747
A/Map position: 17q21-17q22
C/Suprafamily: mouse transcription factor tbx2; T-box homology
F:104-285/Domain: T-box homology <TBX>

Query Match 4.8%; Score 118.5; DB 2; Length 702;
Best Local Similarity 19.3%; Pred. No. 0.63;
Matches 116; Conservative 46; Mismatches 144; Indels 295; Gaps 23;

QY 118 PGVKLRVTEGDEDDGHSEESLHYEGRAVDITTSR-----DRSKYGMUAR 163
Db 95 PKVTLEAKELMDQ-----FKLGTETWITKSGRMFPFKVRVSGLDKKAKYILL-- 144

QY 164 LAVEAGFDWVYVESKAHHCVSVAENS---VAKSGGCGFGSATVHLEGGT----- 212
Db 145 MDIVAADD-----CRYKFNRMWVAGKADPEMPKMYTHPDSPATGQWMAKP 193

QY 213 -----KLVKOLIS-----PGDRVLAADQDGLLYSDFLTFLL----- 242
Db 194 VAFHKLKLTNNISDKHGFTILNSMHKYQPRFHVIRANDILKLPYSTFTYVFPETDFIAV 253

QY 243 -----DRDDCAKVFVIETREPRERLILTAHLL--FVAPHND 279
Db 254 TAYQNDKITQLKIDNNPFAKGRDGTNGR-----REKRKOLLTPSLRLYEHCCKPERD 306

QY 280 SATGPEASSGSGPPS-----GGALGPRALFASRV----- 309
Db 307 GA--ESDASSCDPPPARPPPTSPGAAPSLRLHRAAEKSCAASDDPEPERLSEARA 364

QY 310 -----RPGQRTVVVAERDGD-----RR--- 326
Db 365 PLGRSPAPDSAGFTRITEPARERRCPERGKPEASGSDGPFGLRSLERKEPERARRKDE 424

QY 327 -----LLPAAVHVSILTSEEAAG-----AYAPLTAQQTILINRV 359
Db 425 GRKEAAGKEQGLAPLVVOTDSASPLGAGHLPLGAFSSHLHGQQFFGLGAGQPLFLHP- 483

QY 360 LASCYAVIEHWSWAHRAFPAPRLAHALLAAAPARTDRGDSGGGGRGGRVALTAPG 419
Db 484 -----GQFTMGPCAFSAMGMGH-LLASVA-----GGGNGGG--GGPPTAAGLDAGG 526

QY 420 AADAPGAGATAG--IHWYSQL----- 439
Db 527 LGFAASAATAAPFPFHLSSHMLASOGITPMPTFGGLFPYPTYMAAAAAAASALPATSAA 586

QY 440 -----YQGTWLLDSEALHPLGMVAVKSSXSRGAGGAR 472
Db 587 AAAAAAGSLSRSPFLGARSRLRFPSPYQIPVTIPSTSLATGLA--SEGSKAAGNSR 644
```

QY 473 E 473  
Db 645 E 645

## RESULT 23

E70917  
hypothetical glycine-rich protein Rv1450c - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C/Accession: E70917  
R/Colo: S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Cole, S.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Connor, R.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A/Authors: Seares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
A/Accession: E70917  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1329 <COL>  
A/Cross-references: GB:Z95844; GB:AL123456; NID:G3250713; PID:CA809271.1; PID:G2131046  
A/Experimental source: strain H37RV  
C/Genetics:  
A/Gene: Rv1450c  
C/Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 4.7%; Score 116; DB 2; Length 1329;  
Best Local Similarity 28.3%; Pred. No. 2.2;  
Matches 68; Conservative 18; Mismatches 90; Indels 64; Gaps 12;  
QY 271 LLEFVAPHNDATGEPEASGSG--GPSPGGALGP-----RALFASRVPRPG 312  
Db 3 LVIVAPETVAALDVARIGSSFGAANAAGSTTSVLAAGADEVSAATATLFGSHAREY 62  
QY 313 QRV--VVVAERDGRLLPAHVSVTLSEAGAVAPLTAQGTILINRVLASCYVISEH 370  
Db 63 QAISTQVAAFHDFRQTLGAAGVSY-VSAEATNA-APLATLBNVLNAPTOAL---- 116  
QY 371 SWAHRAFPAPRLAHLL--AALAPARTDRGDSGG-----GDRGGGGGRV 413  
Db 117 -----LGRPLIGDAGAPG-TQAGAGAGGILWNGGAGSGAPGVGGAGGA 164  
QY 414 ALFAP-GADAPAGAGATAGHWYSQLLYQITGWLDSALHPL-GNAVKSXSRGAGGA 471  
Db 165 GLFCTGAGGAGGAGGAGGAG-----GSGGWLGGVGGAGGOSLLGGATGGAGGNA 217

## RESULT 24

E87250  
[protein-pil] uridylyltransferase [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 02-Jun-2003  
C/Accession: E87250  
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, N.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: E87250  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-940 <STO>  
A/Cross-references: GB:AE005673; NID:G13421101; PIDN:AAK22001.1; GSPDB:GN00148  
C/Genetics:  
C/Superfamily: UTP:GlnB (protein Pil) uridylyltransferase

Query Match 4.7%; Score 115.5; DB 2; Length 940;  
Best Local Similarity 23.2%; Pred. No. 1.5;

Matches 82; Conservative 39; Mismatches 155; Indels 77; Gaps 14;  
QY 90 ENTGADRLMTQCKDKLNALAIISVMQWPGVKLRVTGEGDGHSESLHVEGRAVDIT 149  
Db 610 ENPERLRLLLVITVADIRAVPGVWNGWQQLLR-----ELYNATEAVFRGGRSDAA 662  
QY 150 TS-DRDRSKYGMRLARLAVEAGFDWVYYESKAHHCSCVKAENSVAAGSGGCGFFGSAT---- 204  
Db 563 ANVQRQESAEARAAL-----LETPAAKGWVAAMENAYFAFSQDDLF 708  
QY 205 VHLE-----QGGTKLVKDLSPGDR--VLAADQGRLLLYSDFL----- 239  
Db 709 HHAEALARRAAIOGGAAGGVRFGSNAAEVIAAKDR-RGLFADLALAISSLGNVVGAR 767  
QY 240 TELDRDGAKKVYVIE-TREP---RERLLTAAHLLFVAPHNDSATGEPEASGSGPP 294  
Db 768 VFTSRQOQLDVYVQDVTGAPCGCENPRALRLADALEAAGKGDALAVEPRGSGQTRA 827  
QY 295 SGGALGPALF-----ASRVPRGQVYVVAERDGRRLPAPVHSTVLSEEA 341  
Db 828 AAFAPISVITIDNDASNDATVWEASGRDRPG-LLHALAKTLADALSIOAHIDGYGERA 886  
QY 342 AGAVAPLTAOG-----TILINRVLASCYAVIEHSHWAHRAFAP-PRLAHALLA 388  
Db 887 VDAFYVQTTEGGKVTRKLNALKADLLAALAEQNEASAPARGLRPARASVA 939

## RESULT 25

H87311  
hypothetical protein CC0505 [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C/Accession: H87311  
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, N.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: H87311  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-486 <STO>  
A/Cross-references: GB:AE005673; NID:G13421686; PIDN:AAK22492.1; GSPDB:GN00148  
C/Genetics:  
A/Gene: CC0505

Query Match 4.7%; Score 115; DB 2; Length 486;  
Best Local Similarity 21.7%; Pred. No. 0.71;  
Matches 85; Conservative 50; Mismatches 170; Indels 86; Gaps 17;

QY 111 ISVMNQWPGVKLRVTGEGDGHHS-EESLHVEGRAV---DITTSRDRSKYGMRLARLAV 166  
Db 1 METWRATGVLVLTATIGWIAPAHAPAEASAPAKRLIVLSDIENEPPDQTSFVRLLIYAN 60  
QY 167 EAGFDWVYYESKAHHCSCVKAEN--SVAKSGCGPFGSATVHLEQGGTKLVKDLSPGDRV 224  
Db 61 EIDIEALVATTSTHMERGIHPESIRRLVSLYGVQVRP---NLVLHANGYPTAESLSARIRA 117  
QY 225 -----LAADDQGLLYSDFLTFLDRDDGAKVYVIEETPERERLLLT--AAHLLFVAP 276  
Db 118 GQRAYGLAATGPGK-----DTEGSRATIALDSDPPRVWVWVGGAANTLAQAL 166  
QY 277 HNSATGEPEASGSGPPGSGALGPALFASRVPRGQVYVVAERDGD-----RLLPAAV 332  
Db 167 KLEAT-RPAE-----VQRLSKLRVYITISDQDDAGAWIRKTPSLF 208  
QY 333 HSVTLSEAGAVAPLTAQGTI-----LINRVLASCY-----AVIEHSHWAH 374  
Db 209 YIV-----SPGYGAATWGGIFQAVDGLDNTTVSNWLQNIQOGRGPLGAAPVDAIGM 263  
QY 375 RAFAPRLAHALLAALAPARTDRGDSGGGDRGGGRVALTAPGAADAPGAGATAGIHW 434

Db 264 EGTAPFL-NLPTGLA-----DPERDWCWGGRVALYTPNLTDPKGTGGVPI 314  
QY 435 YSOLLYQIGTWLSDSEALH---PLGMVAKSS 462  
Db 315 EPE-TRPIWNAIDTVAPHEPAPFGRAVKVS 344  
RESULT 26  
G70917  
hypothetical glycine-rich protein Rv1452c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 07-Mar-2003  
C:Accession: G70917  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, J.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: G70917  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-741 <COL>  
A:Cross-references: GB:295844; GB:AL123456; NID:G3250713; PIDN:CAB09269.1; PID:G2131044  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: Rv1452c  
C:Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing  
Query Match 4.6%; Score 114; DB 2; Length 741;  
Best Local Similarity 27.9%; Pred. No. 1.5;  
Matches 67; Conservative 18; Mismatches 91; Indels 64; Gaps 12;  
QY 271 LLFVAPHNDGATCEPEASSGS--GPPSGGALGP-----RALFASVRPG 312  
Db 3 LVIVPTETVAASDVARISSGTVANSAAAGSTTVLAAGADEVGAATLFGSHAREY 62  
QY 313 QRV--YVVAERDGRRLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVLASCVAVIEH 370  
Db 63 QAISTQVAAFHREPAQTLAAVGSY-VSAEATNA-APLALTEHNVLNALNAPTQAL---- 116  
QY 371 SWAHRAPFRLAHLL---AALAPATDRGGSGG-----GDRGGGGGRV 413  
Db 117 -----LGRPLTGDGAAGAPG-TGAGGAGGLWNGGAGSGGAPQVGGAGGAA 164  
QY 414 ALTAP-GAADAPGATAGTGHWSQLLYQIGTWLSDSEALHPL-GMAVKSSXSRGAGGGA 471  
Db 165 GLFTGGAGGAGGAGAGGAG-----GSGGMLNGGVGGAGGQSLGCGATGGAGGNA 217  
RESULT 27  
A39065  
homeotic protein EVX2 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 17-Oct-1997  
C:Accession: A39065  
R:D'Esposito, M.; Morelli, F.; Acampora, D.; Migliaccolo, E.; Simeone, A.; Boncinelli, E.  
Genomics 10, 43-50, 1991  
A:Title: EVX2, a human homeobox gene homologous to the even-skipped segmentation gene, F/46-102/Domain: homeobox homology <Hox>  
A:Reference number: A39065; MUID:91257849; PMID:1675198  
A:Accession: A39065  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-333 <DAB>  
C:Genetics:  
A:Gene: GDB:EVX2  
A:Cross-references: GDB:127528; OMIM:142991  
A:Map position: 2q24.3-q31  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F/46-102/Domain: homeobox homology <Hox>

Query Match 4.5%; Score 111; DB 2; Length 333;  
Best Local Similarity 23.2%; Pred. No. 0.87;  
Matches 74; Conservative 27; Mismatches 116; Indels 102; Gaps 12;  
QY 184 SVKAENSVAAGSGCFFPGSATVHLEQGTGLVKDLSGDRVLAAODDGRLLYSDFLTFLD 243  
Db 4 SGSAAGTTTASGS---GLGSLHGGSG-----SGSSAALGGSGGADQVRYRTAFT 53  
QY 244 RDDGAK-KVFF-----VIET-----REPRERLLLTAAH-- 270  
Db 54 REQARLEKEFYRNNYVRPRCELAALNLPETIKWFQNRMMKXQRRLAMSWHPHA 113  
QY 271 ----LLFVAPHNDGATCE-----PEASSGSGPPSGGALGPALFASR 308  
Db 114 DPFVYTMTH-AAATGSLPYFFHSHVPLHYPHVGTAAATAAASGAAAASSPPATS 172  
QY 309 VRQQRVYVVAERDGRRLPAAVHVSITLSEEAAGAY-APLTAQGTILINRVLASCVAVI 367  
Db 173 IRPLDTFRAUSHPHYRPELLCSFRH-----PGLYQAPAAAGLNSAASAAAAAAA 224  
QY 368 EHSWAHRAPFAP-----FRLAHALLAALAPARTDRGGSDGGDRGGGGGR 412  
Db 225 AAASSAAAAGAPPSGGAPSCSLCHSSQAAAAAALGSRGGGGGGGGGGGGG- 283  
QY 413 VALTAPGAADAPGAGATAG 431  
Db 284 -----GGAGAGGG 291  
RESULT 28  
EBBE75  
immediate-early protein IE175 - human herpesvirus 1  
C:Species: human herpesvirus 1  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jun-2000  
C:Accession: A23510  
R:McGeoch, D.J.; Dolan, A.; Donald, S.; Brauer, D.H.K.  
Nucleic Acids Res. 14, 1727-1745, 1986  
A:Title: Complete DNA sequence of the short repeat region in the genome of herpes simpl.  
A:Reference number: A23510; MUID:86148504; PMID:3005980  
A:Molecule type: DNA  
A:Accession: A23510  
A:Residues: 1-1298 <MCG>  
A:Cross-references: GB:X14112; GB:D00317; GB:D00374; GB:S40593; NID:G1944536; PIDN:CAA3.  
C:Comment: This protein acts at the transcriptional regulatory level and is required th  
C:Genetics:  
A:Gene: IE3  
A:Map position: short repeat region (IR-s)  
C:Superfamily: herpesvirus immediate-early  
C:Keywords: DNA binding; early protein; transcription regulation  
Query Match 4.5%; Score 111; DB 1; Length 1298;  
Best Local Similarity 25.8%; Pred. No. 5.1;  
Matches 96; Conservative 34; Mismatches 134; Indels 108; Gaps 20;  
QY 148 ITSDRDRSKY--GMLAR-----LAVTAG-FDWVYYSKAIHCS----- 184  
Db 410 ISCAARNSSSFTTGSVARAFHGLYMAAGRFGWGLAFAAAVAMSRDYRAQKGLTTS 469  
QY 185 -----VKAENSVAAGSGCFFPGSATVHLEQGTGLVKDLSGDR-- 226  
Db 470 LRRVAPLARENAALTGAAGS-PAAGA---DDEGVAVAAGAFGERAVPAGYGAAGILA 525  
QY 227 ADDQGRLLYS-----DPLTFLDRDDGAKK-----VFVITRPREPL---- 264  
Db 526 A--LGRLSAASPAGGDDPDAAAHADADDAGRAQRAQRAVECLAACRGILEALAEGF 583  
QY 265 ---LLTAALLFVAHNDGATCEPEASSG--SGPPSGGALGP-----ALPAS 307  
Db 584 DGLAAMPGLAGARP-----ASPPRPPGAPGAPSPPPHADAPRLRLAWLRELFVRDLVLM 640  
QY 308 RVRPGQRVYVVAERDGRRLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVLASCVAVI 367  
Db 641 RLRGDLRV-----AGGSEAAVAAVRAVSL---VAGALGPALPRDPRLPSSAAAAADLL 691

QY 368 EHS-----WAHRAFPRLAHALLAALAPA--RTDGGDSGGDR--GGGGRVALTAP 418  
DB 692 FDNQSLRLLAAAGAP-DAADALAAAASAAPREGRRKSGFARPPGGGGPRPPKTK 750  
QY 419 GAADAPGAGATA 430  
DB 751 SGADAPGSDARA 762  
RESULT 29  
T35852  
probable UDP-N-acetylmuramoyl-L-alanine ligase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000  
C:Accession: T35852  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.  
submitted to the EMBL Data Library, April 1999  
A:Reference number: Z21591  
A:Accession: T35852  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-462 <SAU>  
A:Cross-references: EMBL:AL049727; PIDN: CAB41553.2; GSPDB: GN00070; SCODEB: SC9B1.07  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCODEB: SC9B1.07  
C:Superfamily: UDP-N-acetylmuramate-alanine ligase  
Query Match 4.5%; Score 110; DB 2; Length 462;  
Best Local Similarity 26.3%; Pred. No. 1.6;  
Matches 47; Conservative 21; Mismatches 71; Indels 40; Gaps 6;  
QY 305 FASRVPRGQVVVVAERDGRLL--LPAVHVSVTL--SEEAAGAYAPLTAQ-----TI 354  
DB 207 FAGKIVGGTTLVIAADHEGARELTRLAGRVRTVYGESEADAVRLSVVPOGLKSEVTV 266  
QY 355 LINVVLASCVAVIEEHSWAHRAFPRLAHALLAALAPARTDGGDSGGGGGRVA 414  
DB 267 VLDGAEITFANVPGRHYANA-----VAALAG-----AA 297  
QY 415 LTAPGAADAPGAGATAGIHWYSQLLYQ--TGTWLLDSEALHPLGMVKSXSGAGGGAR 472  
DB 298 LGVPAEALAPALAAAYTGVRRLQKGEAAGVQVVDVSYAHPTMTADLEAMFAAVGDAR 356  
RESULT 30  
S46458  
transcription factor tbx2 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S46458  
R:Bollag, R.J.; Siegfried, Z.; Cebra-Thomas, J.A.; Garvey, N.; Davison, E.M.; Silver, L.  
Nature Genet. 7, 383-389, 1994  
A:Title: An ancient family of embryonically expressed mouse genes sharing a conserved pr  
A:Reference number: S46458; MUID: 95004605; PMID: 7920656  
A:Accession: S46458  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-701 <BOL>  
A:Cross-references: GB:U15566; NID:G558875; PIDN:AAC52697.1; PID:G558876  
C:Genetics:  
A:Gene: Tbx2  
C:Superfamily: mouse transcription factor tbx2; T-box homology  
F:104-285/Domain: T-box homology <TBX>  
Query Match 4.4%; Score 107.5; DB 1; Length 701;  
Best Local Similarity 20.0%; Pred. No. 4.2;  
Matches 99; Conservative 40; Mismatches 150; Indels 207; Gaps 20;  
QY 118 PGVKLRVTEGDEGHHSSESLHYEGRAVITTSDR-----DRSKYGMLAR 163  
DB 95 PKVTLEAKELWDQ-----FHKLGTEMVITKSGRMFPFPFKVRVSGLDKAKYILL-- 144

QY 164 LAVEAGFDWVYVESKAHHCVSKAENS---VAAKSGGCGPPGSATVHLQGGT----- 212  
DB 145 MDIVAAD-----CRYKFNRSWMVAGKADPEMPKEMYTHPDSPATGEQWMAKP 193  
QY 213 -----KLVKDLS-----PGDVLAAADDOGRLLYSDELTF----- 242  
DB 194 VAFPHKLKLTNNISDKHGFTILNSMHKYQRFHIVANDILKLPYSTFRFYVPPETDFAV 253  
QY 243 -----DRDDCAKKVFYVITREPPRELRLLTAALL--FVAPHND 279  
DB 254 TAYQNDKITQLKIDNPPFAKGFRTGNR-----REKRKQLTLTLRYEHCCKPERD 306  
QY 280 SATGEPEAS-----CSGPPSGALOPRALFASRVPGQVVV-----AERDGERLLP 329  
DB 307 GA--ESDASCDPPAREPPPSAAPSFLRLHRAAEKPGAADSDPEPTGEER--- 361  
QY 330 AAVHVSVTLSEEAAGAYAPLTAQITILINVLASCVAVIEEHSWAHRAFAFFRLAHALLA- 388  
DB 362 -----SAAPLCRSPS-----RDASPAALTEPERSR 386  
QY 389 ALAPARTDRGGSGGGD-----RGGGGRVALTAP----- 418  
DB 387 ERRSPERCSEKTEPGGGDGPFLSLRLEKERPEARXDEGRKDVGECKEPLAPLVVQTOS 446  
QY 419 ----GAADAPGAGATAGIHWY-----SOLLYQITWLLDSEALHPLGM-----A 458  
DB 447 ASPLGAGHLPLGAFSSHLHCQFPGGLGACQPLFLHPGQFAMGPGAFSAMGHHLLASVA 506  
QY 459 VKSXSGAGGAGAREG 474  
DB 507 GGSSSGAGAGGTAAAG 522  
RESULT 31  
EDBE11  
immediate-early protein IE110 - human herpesvirus 1 (strain 17)  
C:Species: human herpesvirus 1  
C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 17-Mar-2000  
C:Accession: A29152  
R:Perry, L.J.; Rixon, F.J.; Everett, R.D.; Frame, M.C.; McGeoch, D.J.  
J. Gen. Virol. 67, 2365-2380, 1986  
A:Title: Characterization of the IE110 gene of herpes simplex virus type 1.  
A:Reference number: A29152; MUID: 87059760; PMID: 3023529  
A:Accession: A29152  
A:Molecule type: DNA  
A:Residues: 1-775 <PER>  
A:Cross-references: GB:X04614; NID:G59832; PIDN:CAA28285.1; PID:G59833  
C:Genetics:  
A:Insertions: 19/3; 242/1  
C:Superfamily: herpesvirus immediate-early protein IE110; RING finger homology  
C:Keywords: DNA binding; early protein; transcription regulation; zinc finger  
F:112-162/Domain: RING finger homology <RNG>  
F:116-156/Region: zinc finger C3HC4 motif  
Query Match 4.3%; Score 107; DB 1; Length 775;  
Best Local Similarity 26.7%; Pred. No. 5.2;  
Matches 68; Conservative 18; Mismatches 89; Indels 80; Gaps 12;  
QY 275 APHNDSATGEPEASGSGPPSGGALG---PRALFASRV-----PQQRVVV-----VAER 321  
DB 371 SPHRPAAAMPGSAPRPGPPASAAASGPARPAAVAPCVRAPPPGPPGPRAPAGAEPAAR 430  
QY 322 DGDRLLPAAVHVSITLSEEAAGAYAPLTAQITILINVLASCVAVIEEHSWAHRAFAFPR 381  
DB 431 PADARRVPOS--HS--SLQAANQEOSLCRARATV-----ARGSGGFGVGGHGPSGAPSPG 484  
QY 382 LAHALLAA-----LAPARTDRG-----GDSGGGDRG 407  
DB 485 AAPLPSAASVQEAAVRPRKRSGQENPSPQSTPPLAPAGAKRAATHPPSDSGPGGRG 544  
QY 408 GGGGRVALTAPGAADAPGAGATAGIHWYSOLLYQITWLLDSEALHPLGM-----AVKS 461

Db 545 QGPGTTLTS-SASASSSSASS-----SSAPTPGASASAAAGASS 585

Qy 462 SXSRGAGG--GAREG 474  
| | | | |  
Db 586 SASASSGGAGALGG 600

RESULT 32  
E70554  
hypothetical protein Rv1148c - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: E70554  
R: Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
A: Authors: Squares, R.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrall, B. G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: E70554  
A: Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-482 <COL>  
A: Cross-references: GB: 235584; GB: AL123456; NID: 93261774; PIDN: CAB09036.1; PID: g2117186  
A: Experimental source: strain H37Rv  
C: Geneticals:  
A: Gene: Rv1148c  
C: Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1945

Query Match 4.3%; Score 106.5; DB 2; Length 482;  
Best Local Similarity 22.0%; Pred. No. 3.1;  
Matches 87; Conservative 40; Mismatches 151; Indels 117; Gaps 16;

Qy 49 PNVAEKLTSAGRYEGKISNSER-----FKELTPNPNPDIIFKDEENTGADRLMTQRCK 103  
| | | | |  
Db 137 PLPAQLTATAAQREGKIGREHIKEIQAFKELSAV--DLGIREAAEAQELATSRP 194  
| | | | |  
Qy 104 DKLNALAISVMNPGVKLRVTGEWDGHHSESLHYEGRVAVDITTSDDRS- KYGMLA 162  
| | | | |  
Db 195 DHLHGLATQLMD-----W-----LHPDG-----NFSQERAKRGITM 227  
| | | | |

Qy 163 RLVAEAGFDWVYVESKAHIGSVKAEKNSVAKSGCGPFGSATVHLEQGGTKLVKDLSPGD 222  
| | | | |  
Db 228 GKQFDMGSRISGLLTTELRAITAEVLAKLAAPACNPDDQ-----TFLVDIDPDAD 279  
| | | | |  
Qy 223 RVLA-ADDQGRLLYSDFLFLDRDDGAKKVFYVIETREPRERLLTLTAHL-----LFVAP 276  
| | | | |  
Db 280 AVRRDTRSAQRNHDALFLAAL-----RGLLAGSGELGQHKGLPVTI 319  
| | | | |

Qy 277 HNDSATGEPEASSSGSPSGGA-----LGPRLAFASR--VRP 311  
| | | | |  
Db 320 VVSTTLKELEATGKVTGGGSRVPMGDLIRMAASHANHYLALFGAKPLALYHTKRLASP 379  
| | | | |

Qy 312 GQRYVYVAERDG-----DRELLPAVHVSFTLSEAGAVAPLTAQGTILINRVLASCYA 365  
| | | | |  
Db 380 AQRIMLYAKDGRGCRPGCDAPVHSEVHVT-----PWTTHRTDINDTLACGP 429  
| | | | |

Qy 366 --VIEEHSWAHRAFPFLAHALLAALAPARTDRG 398  
| | | | |  
Db 430 DNRILVEKGWKTNR-----KNARGDTEWLPFPHLDHG 459  
| | | | |

RESULT 33  
JCT583  
basic helix-loop-helix protein, DEC2 - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JCT583  
R: Fujimoto, K.; Shen, M.; Noshita, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.  
Biochem. Biophys. Res. Commun. 280, 164-171, 2001  
A: Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loop  
A: Reference number: JCT583; MUID: 21092582; PMID: 11162494

A: Accession: JCT583  
A: Molecule type: mRNA  
A: Residues: 1-482 <FUJ>  
A: Cross-references: DDBJ: AB044088  
C: Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix  
C: Geneticals:  
A: Gene: dec2  
A: Map position: 12p11.23-p12.1  
C: Keywords: transcription factor  
C: Region: highly conserved #status predicted  
F130-173/Domain: Orange #status predicted <ORA>  
F1286-411/Region: alanine and glycine-rich #status predicted

Query Match 4.3%; Score 106; DB 2; Length 482;  
Best Local Similarity 24.7%; Pred. No. 3.4;  
Matches 77; Conservative 21; Mismatches 94; Indels 120; Gaps 16;

Qy 204 TVHLEOGGTKLVKDLSPGDRVLADDDGRL--LYSDFLTFLDRDDGAKKVFYVI---ETR 258  
| | | | |  
Db 100 TALTEQQHQKIIA-LONGERSLKPIQSDLDFAHSGFTC-----AKEVLOYLSRFESW 152  
| | | | |

Qy 259 EPRE-RLLLTAHLLFVA-----PHNDSATGEPEASSSGSPPS--GGALGPRLAFASVRP 311  
| | | | |  
Db 153 TPPEPCVOLINHLHVAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAP-----CLERA 207  
| | | | |

Qy 312 GQRYVYVAERDG-----DRELLPAVHVSFTLSEAGAVAPLTAQGTILINRVLASCYA 365  
| | | | |  
Db 208 SOKLEPLAYCVFVIQRTOPSAELAAENDTDTDSGVGGEAEARPDREKKGAGASRVTIKQ 267  
| | | | |

Qy 340 EAAGAYAPLTAQGTILINRVLASCYAVIEHSWAHRAFPFLAHALLAALAPARTDRGG 399  
| | | | |  
Db 268 EPPGEDSP-----APKRM-----KLDSSRG 287  
| | | | |

Qy 400 DSGGDRGGGGRVALTAPGAADAPGACATAGIHWYSQLLYQIGTWLWLDSEALHPLGNV 459  
| | | | |  
Db 288 GSGGGPFGG-----AAAAAALLGDPAAAAAL-----LRPDAALL 323  
| | | | |

Qy 460 KSSXSRGAGGGA 471  
| | | | |  
Db 324 SSLVAFGGGGGA 335  
| | | | |

RESULT 34  
T09084  
phosphatidylinositol 3-kinase - Chlamydomonas reinhardtii (fragment)  
C:Species: Chlamydomonas reinhardtii  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T09084  
R: Molendijk, A. J.; Irvine, R. F.  
Plant Mol. Biol. 37, 53-66, 1998  
A: Title: Inositolide signalling in Chlamydomonas: Characterization of a phosphatidylinosi  
A: Reference number: Z16411; MUID: 98281574; PMID: 9620264  
A: Accession: T09084  
A: Status: Preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-490 <MOL>  
A: Cross-references: EMBL: U97663; NID: g2109290; PIDN: AAC50018.1; PID: g2109291  
A: Experimental source: strain cw-15  
C: Geneticals:  
A: Introns: 265/3; 331/3; 370/3; 455/1; 481/3

Query Match 4.3%; Score 105.5; DB 2; Length 490;  
Best Local Similarity 25.1%; Pred. No. 3.7;  
Matches 76; Conservative 23; Mismatches 123; Indels 81; Gaps 15;

Qy 191 VAAKSGCGPFGSATVHLEQGGTKLVKDLSPGDRVLADDDQGRLLYSDFLTL----DRDD 246  
| | | | |  
Db 234 LAAGGGG---GGG-----SPGDSTAR-----WDEWLTFCVKYRDLPP 271  
| | | | |

Qy 247 GAKKVFYVIETREPRERLLTLTAHLFVAPHNDSATGEPEASSSGSPSGGALGPRLAFA 306  
| | | | |  
Db 272 DAQLVLLVLAEGREALVGVSSVTLFSGKRLTKGPORLAVWEGAP-----PCTQPP 325  
| | | | |

```

C:Species: Streptomyces fradiae
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T14070
R:Hosted, T.J.; Baltz, R.H.
  submitted to the EMBL Data Library, July 1997
A:Reference number: Z17868
A:Accession: T14070
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1324 <HOS>
A:Cross-references: EMBL:AF016696; NID:g2738764; PID:g2738765; PIDN:AAC01566.1
C:Genetics:
A:Note: cps-1
C:Superfamily: Mycobacterium tuberculosis mbtE protein; acetate-CoA ligase homology; acy
C:Keywords: carrier protein; phosphoprotein; phosphoprotein
F:188-642/Domain: acetate-CoA ligase homology <AC1>
F:659-727/Domain: acyl carrier protein homology <ACP>
F:691/Binding site: phosphoprotein (Ser) (covalent) #status predicted

Query Match      4.3%; Score 105.5; DB 2; Length 1324;
Best Local Similarity 22.2%; Pred. No. 14;
Matches 108; Conservative 52; Mismatches 188; Indels 139; Gaps 21;

Qy      52 AKETIGASG-RVEGKISRNSRFK-----ELTPNYNPDIIFKDEENTGADRLMT 99
Db      46 AVSSAGHAGIRIEGGSGRTTNHFP:SLYALPGALRLRDLDRHPDAV-----DDVT 96
Qy      100 QRCKDKL-----NALATSVNQWPGVKLRVTGWDGDGHSE 136
Db      97 RHAADLLERALTAVHSAPATPTAALAAATPATRAAAPRAAGFGAPATIVDAFEARVRAFP 156
Qy      137 ESLHYEGRAVDITTSDRDSKYGMRLARLAVEAGFDWVYVESKAHHCVSKAENSVAA---- 193
Db      157 EAPAVLAGGEELTYAELD-ARANRLARLLRLRG---VGPESRVALTYSNAMLPAVLGI 212
Qy      194 -KSGGCF-FGSATVHLEQCGTKLVKDLSPGDRVLAADDQGRLLYSDFLFLFLDRDDGAKV 251

```

```

Db      213 LKAGGCVFVGATLPRER--AARILRETAVICLLTDPDAEAARTTRTAPTGGDDRDENA--- 268
Qy      252 FYVIEIIEPR--ERILLTAAHLLFV--APHNDSATGEP-----EASSGSGPPSGG 297
Db      269 -----FGVERVVLITGALLAAFPDAPPTDAERAGPLLPGLHAYLLHTSSSGPKGV 320
Qy      298 ALGPRALFASVRPQGQVVVVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQGTI--- 354
Db      321 AVEHAQVALLSWAGTGV-----GADRLHRTVASTSESFVSVFDTLVPLLTGGRIEV 374
Qy      355 -----LINRVLASCYAVIEE-----HSM--AHRAPAPRLAHALLAALA 391
Db      375 ENTILAVADRTGGBPSLINAVPSALQALLERGEPLAVHTFLCAGEPF--PAPLARSULRAAPP 433
Qy      392 PAR-----TDRGDSGGGDRGGGGRVALTAPGAADAP-----GA 426
Db      434 RARVANIYGPTETTVFVTAHFLDGTDDGAPPVG--RPLFGVRVHILDPLMLRPVFDGVVGE 491
Qy      427 GATAGIH 433
Db      492 LYLAGEH 498

RESULT 37
T36248
CDA peptide synthetase I - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 01-Dec-2000
C:Accession: T36248
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21602
A:Accession: T36248
A>Status: preliminary, translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7463 <SAU>

```

A;Cross-references: EMBL:AL035640; PIDN: CAB3851.8.1; GSPDB: GNO0070; SCODEB: SCE63.03c	
A;Experimental source: strain A3(2)	
C;Genetics:	
A;Gene: cdaPSI; SCODEB: SCE63.03c	
C;Superfamily: acetate-CoA ligase homology; acyl carrier protein homology	
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein	
F;516-1074/Domain: acetate-CoA ligase homology #status atypical <ACLI>	
F;1090-1158/Domain: acyl carrier protein homology <ACP1>	
F;1715-2184/Domain: acetate-CoA ligase homology <ACLI>	
F;2200-2268/Domain: acyl carrier protein homology <ACP2>	
F;2804-3249/Domain: acetate-CoA ligase homology <ACLI>	
F;3265-3332/Domain: acyl carrier protein homology <ACP3>	
F;4323-4746/Domain: acetate-CoA ligase homology <ACLI>	
F;4762-4830/Domain: acyl carrier protein homology <ACP4>	
F;5363-5786/Domain: acetate-CoA ligase homology <ACLI>	
F;5802-5870/Domain: acyl carrier protein homology <ACP5>	
F;6401-6868/Domain: acetate-CoA ligase homology <ACLI>	
F;6884-6951/Domain: acyl carrier protein homology <ACP6>	
F;1122,2232,3297,4794,5834,6916/Binding site: phosphopantetheine (Ser) (covalent) #status	
Query Match 4.3%; Score 105.5; DB 2; Length 7463;	
Best Local Similarity 19.6%; Pred. No. 1.3e+02;	
Matches 116; Conservative 34; Mismatches 170; Indels 279; Gaps 25;	
QY 118 PGVKLRVTEGWDSDGHH-----SESLHYEGRAVDITSD--RDRSK 157	
DB 450 FDERERVLDGNDTAHEVPETTLPELFAAPARTPGHEALVYEGTSLTYAELDARAERLA 509	
QY 158 YGMLARLAVEAGFDWYVESKAHHCVSVAENSAKSGGCF-----199	
DB 510 GALTARGAGPERFVAVAVERSAELVALLA-----VLKSGNAYVPDPGYPADRIAHLRD 565	
QY 200 -----PGSATVHL-----207	
DB 566 AGAMLVLTTRDTAERLEGGDTGPRLLLDPEAAAGTTAAGAPAPGTLFRALPAPGEPAYVI 625	
QY 208 -EQGTYLVK-----DLSPGDRVLAADD-----OG 231	
DB 626 YTSGSTGRPGVIVISHRAIYNRLAWMDTYGLEPDRVLQKTPSGFDVSWFEFVLVQG 685	
QY 232 RLL-----YSD--FLTFLDRDDGAKKVFV-----IETREPRRL--LTAHLLF-- 273	
DB 686 ATLVVAPGGHTDPAYLAGTVRRSGVTLHFVPSMLDVFLEPAAALGGATPVRVFC 745	
QY 274 -----VAPNDSATGEPAS-----SGSGP-----293	
DB 746 GEALPSELRAFRFRAVSDVPLHN--LYGFTAAVDVTWVPCAEOTGDPGVFPIGRPVNTRM 803	
QY 294 -----PSG-----GALGPALFASRV-----RPGQVYVVAE-- 320	
DB 804 YVDAALRPVPAGVPGELYIAGVQLARGYLGRPALSAERTTADPHGAPGSRMYRTGLAR 863	
QY 321 --RGDRRLLPAAVHSTVL-----SEEAAGAYAPLTAQGTILINRVLASCVAVIEEH 370	
DB 864 WNHGSLDYLGRADHQVKGRLGRIELGEIEAALVRQEIQAQAVLR-----EDR 913	
QY 371 SWARAPFA--PFLALHALAALAPARTDRG-----GDSGGDRGGGGGRVALT 416	
DB 914 PGDQLVAYTPARDADTLTGPPAEAGTHPGFAPDTPATGPAAGTDSGPGSG-----T 969	
QY 417 APGAADAPGATAGIHWYSOLLYQIGTWLLDLSBALHPL-----GMAVKSXSRGAGGARE 473	
DB 970 GSGTSGTGSGARPG-----PDGTATHTVAGAGPAAAGETAAGADACTGT 1014	
QY 474 G 474	
DB 1015 G 1015	
RESULT 38	
T33110	
C;Species: Caenorhabditis elegans	
C;Superfamily: glycerol-3-phosphate dehydrogenase (aerobic)	
C;Keywords: oxidoreductase	
Query Match 4.2%; Score 104.5; DB 2; Length 510;	
Best Local Similarity 23.5%; Pred. No. 4.7;	
Matches 91; Conservative 35; Mismatches 123; Indels 139; Gaps 20;	
QY 103 KDKLNALAI SVNQWPGVKLRVTGDEDEGHHSLSHYEGRAVDITSDRDRSKYGM 162	
DB 125 REKLPSARACVHRQOP-LKAEIRRGFE-----YSDCAVDD-----A 159	
QY 163 RLAVEAGFDWYVESKAHHCVSVAENSAKSGGCFPGSAT-----204	
DB 160 RLVLNLAIS--AREHGAHVHTAPASAPVAARDSTGCTCTWSAATAACIRSRALVNAGFW 217	
QY 205 -----VHLEQGGTKLVKDLSPGDRVLAADDGGRLL-----YSDFLTFI-- 242	
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000	
C;Accession: T33110	
R;Tin-Wollam, A.; Fronick, W.	
submitted to the EMBL Data Library, May 1998	
A;Description: The sequence of C. elegans cosmid C18H7.	
A;Reference number: 221284	
A;Accession: T33110	
A;Status: preliminary; translated from G5/EMBL/DBJ	
A;Molecule type: DNA	
A;Residues: 1-460 <TIN>	
A;Cross-references: EMBL:AF067607; PIDN: AAC17641.1; GSPDB: GNO0022; CESP: C18H7.3	
A;Experimental source: strain Bristol N2; clone C18H7	
C;Genetics:	
A;Gene: CESP: C18H7.3	
A;Map position: 4	
A;Introns: 84/1	
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8	
Query Match 4.3%; Score 105; DB 2; Length 460;	
Best Local Similarity 28.9%; Pred. No. 3.8;	
Matches 67; Conservative 9; Mismatches 100; Indels 56; Gaps 11;	
QY 275 APNDSATGEP-----EASSGSGP-PSG--GALGPALFASVRPQRYVYVVAERD- 323	
DB 207 APGEPGAPGPPGPPNGEAGAGSGPGEPPGPPGNGKDGAGAPG-----KDG 258	
QY 324 -----DRLLPAAVHSTVL-----SEEAAGAYAPLTAQGTILINRVLASCVAVIEEH 370	
DB 259 AGADGAPGTDAAFCPCPPRSAAALGAGGGAEPAGA-APEAA-----AAAPAAPEAA 308	
QY 371 SWAH-----RAFAPFLAHALLAALAPARTDRGDSGGDRGGGGGRVALTAFGAA 421	
DB 309 PAEGGNGGAPGAPGAPDAAAAPAAPEAAPEAAPEAGGGAEPAGAAPDAAAAPAA 368	
QY 422 DAPGAGATAGTHWYSOLLYQIGTWLLDSEALHPLGMVKSS--XSRGAGGGA 471	
DB 369 PAEMAPAAEGAGGGAEE---PAGAAPDAAAAPAAPEAAPEAAPEAGGGA 417	
RESULT 39	
A55207	
glycerol-3-phosphate dehydrogenase glpD (EC 1.1.1.-) - Pseudomonas aeruginosa	
C;Species: Pseudomonas aeruginosa	
C;Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 11-Jun-1999	
C;Accession: A55207	
R;Schweizer, H.P.; Po, C.	
J. Bacteriol. 176; 2184-2193, 1994	
A;Title: Cloning and nucleotide sequence of the glpD gene encoding sn-glycerol-3-phosph	
A;Reference number: A55207; MUID: 94209216; PMID: 8157588	
A;Accession: A55207	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-510 <SCH>	
A;Cross-references: GB:L06231; NID: g450377; PIDN: AAA81584.1; PID: g450378	
C;Genetics:	
A;Gene: glpD	
C;Superfamily: glycerol-3-phosphate dehydrogenase (aerobic)	
C;Keywords: oxidoreductase	
Query Match 4.2%; Score 104.5; DB 2; Length 510;	
Best Local Similarity 23.5%; Pred. No. 4.7;	
Matches 91; Conservative 35; Mismatches 123; Indels 139; Gaps 20;	
QY 103 KDKLNALAI SVNQWPGVKLRVTGDEDEGHHSLSHYEGRAVDITSDRDRSKYGM 162	
DB 125 REKLPSARACVHRQOP-LKAEIRRGFE-----YSDCAVDD-----A 159	
QY 163 RLAVEAGFDWYVESKAHHCVSVAENSAKSGGCFPGSAT-----204	
DB 160 RLVLNLAIS--AREHGAHVHTAPASAPVAARDSTGCTCTWSAATAACIRSRALVNAGFW 217	
QY 205 -----VHLEQGGTKLVKDLSPGDRVLAADDGGRLL-----YSDFLTFI-- 242	

Db 218 VARFIQDULKQSPYGIIRLIQSHIIVPKLVEGEHAYILQNEDDRIVFAIYDLRFTWIG 277  
QY 243 --DRD---DGAK-----KVFYVITREPRERLLLTAHLLF-----VAPHNDSATGEPEA 287  
Db 278 TTDREVGQDPKAVISEETALLQVNAHFQQLAAADILHSPAGVPLCDSDSDPSA 337  
QY 288 -----SSSGSPPS-----GGALGP-RALFASVRPQGVVVAERDGRLLPAAV 332  
Db 338 ITRDYTLISLISAGNBPPLLSVFGKLTYYRKLAEALTLQLOPFA-----NLGPWT 389  
QY 333 HSVTLSEEAAGAVAPLTA---QGTILINRVLASCVAVIEEHSWAHRAFAFAPRLAH----- 384  
Db 390 -----AXAPLPGEQWQSVETALTEQLANRYA-----WLDRELAPALGAHYCTR 432  
QY 385 --ALLAALAPARTDRGDSGGDRGGG 410  
Db 433 VWRLLDGV-----NGEADLGEHLGG 453

RESULT 40  
S25077  
monensin polyketide beta-ketoacyl synthase (EC 2.3.1.-) chain 2 - Streptomyces cinnamonome  
C;Species: Streptomyces cinnamonensis  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: S25077; S18168  
R;Arrowsmith, T.J.; Maipartida, F.; Sherman, D.H.; Birch, A.; Hopwood, D.A.; Robinson, J.  
Mol. Gen. Genet. 234, 254-264, 1992  
A;Title: Characterisation of acti-homologous DNA encoding polyketide synthase genes from  
A;Reference number: S25076; MUID:92374994; PMID:1508151  
A;Accession: S25077  
A;Molecule type: DNA  
A;Residues: 1-402 <ARR>  
A;Cross-references: EMBL:Z11511; NID:G46799; PIDN:CAA77597.1; PID:G46801  
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I; 3-oxoacyl-[acyl-carrier-pro  
C;Keywords: acyltransferase  
F;22-398/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>

Query Match 4.2%; Score 104; DB 1; Length 402;  
Best Local Similarity 24.1%; Pred. No. 3.7;  
Matches 95; Conservative 33; Mismatches 140; Indels 126; Gaps 21;

QY 140 HYGRAVDITSDRSKYGMLARLAVEAGFDWVYESKAHICSVKAENSV--AAKSGG 197  
Db 60 HLPGLVPQT-----VTRLALAA-DWALADAGVEVAAPPLDNGVVTASHAGG 109  
QY 198 CFFGSATVHLEQGGTKLVKDSPLGDRVLAADDQGRLLYSDFLFLDRDDCAKVFY----- 253  
Db 110 -----FEFGDELQKLGGQPVLSA-----YQSFAMFYAVNSGQISIRHGMKG 153  
QY 254 -----VIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSPSPGALGPR 302  
Db 154 PSGVVSQDQAGGLDALAQARRLVKGTPLIV-----CGAVEPRSPAGGSPSPAGG-- 205  
QY 303 ALFASRVPRQGVVVAERDGRLLPAAVHSVTLSEEAAGAVAPLTAQGTILINRVLAS 362  
Db 206 ---MSDSDENRAVLPEDRUG-RGYVPGGCGGVVPLERAEA-AP--ARGA----- 249  
QY 363 CYAVIEHSHWAHRAFAF-----RLAHLAALAPARTDRG-----GDSGG--GDRGG 408  
Db 250 --EYVGEAGPLARLPAPSHGRGSTRAHAIRTAALDDAGTAFDIRRVFADGGGGRYPNDRAE 307  
QY 409 -----GGRVALTAP-----GAADAPAGA-----TAGI-----HWY 435  
Db 308 ABATSEVFGPRVPTCFRTWGTGLHSGAAPLDVACALLAMRAGVTPPTVHIDPCPEYDL 367  
QY 436 SLLYQGTWLLDSEALHPLGMVKSXSRGAGG 469  
Db 368 DLVLYQV-----RPAALRTALGGARGHG 391

RESULT 41  
A40253  
acidic nuclear protein SPT5 - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein YM9571.08; protein YML010W  
C;Species: Saccharomyces cerevisiae  
C;Date: 06-Dec-1991 #sequence\_revision 06-Dec-1991 #text\_change 29-Oct-1999  
C;Accession: A40253; S55109  
R;Swanson, M.S.; Malone, E.A.; Winston, F.  
Mol. Cell. Biol. 11, 3009-3019, 1991  
A;Title: SPT5, an essential gene important for normal transcription in Saccharomyces cer  
A;Reference number: A40253; MUID:91246167; PMID:1840633  
A;Accession: A40253  
A;Molecule type: DNA  
A;Residues: 1-1063 <SWA>  
A;Cross-references: GB:M62882; NID:G172579; PIDN:AAA35085.1; PID:G172680  
R;Gentles, S.; Bowman, S.  
submitted to the EMBL Data Library, June 1995  
A;Reference number: S55102  
A;Accession: S55109  
A;Molecule type: DNA  
A;Residues: 1-1063 <GEN>  
A;Cross-references: EMBL:Z49810; NID:G854472; PIDN:CAA89942.1; PID:G854480; MIPS:YML010W  
A;Experimental source: strain AB972  
C;Genetics:  
A;Gene: SGD:SPT5  
A;Cross-references: SGD:S0004470; MIPS:YML010W  
A;Map position: 13L  
C;Keywords: nucleus; tandem repeat  
F;149-216/Region: aspartic acid/glutamic acid-rich  
F;931-1063/Region: 6-residue repeats

Query Match 4.2%; Score 104; DB 2; Length 1063;  
Best Local Similarity 21.6%; Pred. No. 13;  
Matches 119; Conservative 61; Mismatches 203; Indels 168; Gaps 29;

QY 45 KQFIPNVAEKLGLASGRYEG-----KISRNSERFKLTPTNPNP--IIFKDEINTGADRL 97  
Db 496 KNIQPTVEE--LARFGSKEGAVDLTSVSQIKKAQKVTFFQGDRIEVLNGQRSGKGI 553  
QY 98 MTQCKD-----KLNALA-----ISVMNQ--WPGVKLRVTEGWDGCHSESL----- 139  
Db 554 VTRTKDOIATIKLNGFTTPEFFISTURKIFEPGDHVTING-----EHQGDAGLVLMVEQ 609  
QY 140 -----HYGRAVDITSDRSKYGMLARLAVEAGFDWVYESKAHICSVKAENSV- 191  
Db 610 QGVTFMSTQTSREVTITANNLSKS-IDTTATSEYALHDLIVLSAK-NVACIIQAGHDIF 667  
QY 192 -----AAKSGCFCPSATVHLEQGGTKLVKDSLPGDRVLAADD-----QGRLLYSD 237  
Db 668 KVIDETGVSTITKGSILSKINTARARVSSVDANGNEIKIGDTIVEKVGSRREGGVLYIQ 727  
QY 238 FLTFLDRDDGAKVYVYVETREPRERLLLTAHLLFVAPHN-----DSAT 282  
Db 728 -----TQQIFVV-----SKKIVENAGVVFVNPSNVEAVASKDNMLSKMDLSK 770  
QY 283 GEPASSGSGSPS-----GG---ALGPRALFASRVPRQGVVVAERDGRLL 328  
Db 771 MNPEIISKMGPPSKTFFQIQSRGGREVALGKTVRIRSAGYKGO-LGIVDVGNDKATV 829  
QY 329 PAAVHS-----VTLSEEAAGAVAPLTAQGTI---LINRVLASCVAVIEEHSWAHRAFAFPR 381  
Db 830 --ELHSKNKHITIDKHKLTYNREGGEGITVDELVNR---RGRVPQARMGSPSVSAPRN 883  
QY 382 LAHALAALA-----PARTDRGDSGG-----DRGGGGG 411  
Db 884 MATGGIAAGAAATSGLSGGMTPOWSFDGKTPAVNAHGGSGGGVSSWGGASTWGGQG 943  
QY 412 RVALTAPCAADAPAGAGATAG:HWYSQLLYQIGTWLLDSEALHPLGMVAV---KSS-----XS 464  
Db 944 NGGASAWGGA---GGGASA---WGGQGTGATSTW-----GGASAWGNKSSWGGAS 987  
QY 465 RGAGGGAREGA 475  
Db 988 TWASGGESNGA 998

## RESULT 42

T30342  
C:Species: Yersinia enterocolitica  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T30342  
R: Palludat, C.; Rakin, A.; Jacobi, C.; Schubert, S.; Heesemann, J.  
J. Bacteriol. 180, 538-546, 1998  
A: Title: The yersinabactin biosynthetic gene cluster of Yersinia enterocolitica: organization and function  
A: Reference number: 220833; PMID: 98117033; PMID: 9457855  
A: Accession: T30342  
A: Status: preliminary; translated from GB/EMBL/DDBJ  
A: Molecule type: DNA  
A: Residues: 1-3161 <PEL>  
A: Cross-references: EMBL:Y12527; NID:el228641; PID:el228642; PIDN:CAA73127.1  
C: Genetics:  
A: Gene: irp1  
C: Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protein  
C: Keywords: carrier protein; phosphopantetheine; phosphoprotein  
F: 47-447/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>  
F: 1816-1887/Domain: acyl carrier protein homology <ACP1>  
F: 2825-2892/Domain: acyl carrier protein homology <ACP2>  
F: 2856/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 4.2%; Score 104; DB 2; Length 3161;  
Best Local Similarity 20.9%; Pred. No. 54;  
Matches 106; Conservative 51; Mismatches 142; Indels 208; Gaps 24;  
QY 105 KLNALATSVNORGVK-LAVTE-----GWDED--GHSSESLHYEGRVDTIT 150  
DB 137 KTGWFASRMSTYPGREALNVTEVAQVKGLOSLMGNDKQVIATRAAYKLNHGPALSQT 196  
QY 151 SDRDSRYGMLARLAVEAGFDWVYVESKAHIC-SVKAENSVAASKGG---CPFGSATVH 206  
DB 197 A-----CSSSLVA-----VHLACESLRAGESDMAVAGGVALSFPQAGYR 236  
QY 207 LEQG-----GTLVKDLSPGDRVLADDQGRLLYSDFLTF 241  
DB 237 YQPGMIFSPDGHCRPFDAEGTWAGNGLCVLRRLR--DALLSGDP-----IISVILSS 290  
QY 242 LDRDGAQKVFYVETREPRERLLTAHLLFVAPHNDSATGEPFASSGSGPSPGGALGP 301  
DB 291 AVNNDGNRKVGYTAPSVAGQAVIEEA--LMLAIDDRQGVIE-THGTGLPLGDAIEI 346  
QY 302 RALFASVRPQQRVYVAERDGRRLPAAVHS----- 334  
DB 347 EAL-----RNVY--APRPDQRCALGSYKSNMGLHDTAAGTAGLLKTVLAVSRGQIP 396  
QY 335 -----VTLSEEAAGA 344  
DB 397 PLLNHTPNPAKLEESPTIPVSAQWQDMRYAGVYSSFGIGGTNCHMIVASLPDALNA 456  
QY 345 YAPLTAQG-----TVLASYA-VIEHS-----WAHRAPAPFRL 382  
DB 457 RLPNTDSGRKSTALLSAAASDALLRLATDYAGALRENTDASDLAFTALHARLDLPFRL 516  
QY 383 AHALLAALAPARTDRGDSGG-----GDRGGGGGRVALTAPGAADPAGAGATAGLHW--YS 436  
DB 517 AAPLNRETAALSDWAGKSGALVYSGHAGSKQVWLF-----GQGS-----HWRTMG 565  
QY 437 QLLYQIGTWLLD-----SEALHP 454  
DB 566 QTMVQHSYAFADMLDRCFACSEMLTP 592

## RESULT 43

S33121  
N: Alternate names: CCAAT displacement protein; homeotic protein Clox  
C: Species: Canis lupus familiaris (dog)  
C: Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Aug-1999  
C: Accession: S33121  
R: Andres, V.; Nadal-Ginard, B.; Mahdavi, V.

Development 116, 321-334, 1992  
A: Title: Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding protein  
A: Reference number: S33121; PMID: 93161940; PMID: 11363085  
A: Accession: S33121  
A: Status: preliminary  
A: Molecule type: mRNA  
A: Residues: 1-975 <AND>  
A: Cross-references: EMBL:X69017; NID:g2201; PIDN:CAA48782.1; PID:g2202  
C: Superfamily: homeotic protein CDP; cut repeat homology; homeobox homology  
C: Keywords: DNA binding; duplication; homeobox; nucleus; transcription regulation  
F: 1-67/Domain: cut repeat homology (fragment) <CU1>  
F: 383-455/Domain: cut repeat homology <CU2>  
F: 566-638/Domain: cut repeat homology <CU3>  
F: 685-741/Domain: homeobox homology <HOX>

Query Match 4.2%; Score 103.5; DB 2; Length 975;  
Best Local Similarity 23.8%; Pred. No. 13;  
Matches 71; Conservative 21; Mismatches 97; Indels 109; Gaps 10;  
QY 166 VEAGFDWVYVESKAHICSVKAENSVAASKGGCGPFGSATVHLEGGTKLVKDLSPGDRVL 225  
DB 635 VEKLMDMKMEKKAY----MKRRHSSVSDSQCPSPSVGIDYSQSGSPQOHLKPRV 690  
QY 226 AADDOGRLL-----YSDFLTFLD-EDDGAKKVFYVETREPRERLLTAAH 270  
DB 691 LAPEEKALKRAYOQKPYSPKTIIEELATQNLKSTVINWFHNYRSIRRE----- 742  
QY 271 LLFV-----APHNDSATGEPFASSGSGPSPGG-----ALGPRALPASVR 310  
DB 743 -LFIEIQAGSQGAGARHSPSA-----RSSGAAPSSSEDCDGVAAEGPGAADA----- 792  
QY 311 PQQRVYVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAQGTILINRVLASCYAVIEEH 370  
DB 793 -----EESAPAAAKSQGGP----- 807  
QY 371 SMAHRAPAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGADAFAGA 428  
DB 808 -AEAAPAEEREEAPRAEKRSRPRGPGPGRRGGG-----PAGGAPAAPAAA 858  
RESULT 44  
E70661  
probable PE protein - Mycobacterium tuberculosis (strain H37Rv)  
C: Species: Mycobacterium tuberculosis  
C: Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C: Accession: E70661  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; PMID: 98295987; PMID: 9634230  
A: Accession: E70661  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-413 <COL>  
A: Cross-references: GB:Z83860; GB:AL123456; NID:g3261681; PIDN:CA806157.1; PID:g1781247  
A: Experimental source: strain H37Rv  
C: Genetics:  
A: Gene: PE  
C: Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 4.2%; Score 102.5; DB 2; Length 413;  
Best Local Similarity 25.5%; Pred. No. 5;  
Matches 51; Conservative 17; Mismatches 77; Indels 55; Gaps 6;  
QY 297 ASSGGSPSGGAL-----GPRALPASVRPQQRVYVAERDGRRLPAAVHSVTL 337  
DB 28 ANAAAAPTAGVLAAGGDDVSAGIALFGARQAQYQAISQAALPHDR-----FQVI 79  
QY 338 SEEAAGAYAPLTAQGTILINRVLASCYAVIEHSHWAHRAFPRLAHALLAALAPARTDR 397

Db 80 LOEGAAAYAMAEANAL-----PLOKAOQVSVSELAQDRT-- 113  
QY 398 GDSGGGDRGGGGGRVALTPGAADAPGAGATAGIHWTYSQLLYQITGTWLLDSEALHP--- 454  
Db 114 GTGTGTGQSEAGG-----FGVGAGGKWDGGPIGNQGVQGEHGAGQLGSTDGNGPVA 167  
QY 455 -----LGMVKSXSXSRGAGGA 471  
Db 168 GAAGSGVSASRSGATGAA 187

RESULT 45  
T27975  
hypothetical protein Rv0341 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: F70573  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Hollroyd, S.; Rajandream, V.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: F70573  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-479 <COL>  
A: Cross-references: GB:295324; GB: A1123456; NID: 93261760; PIDN: CAB08573.1; PID: G2094820  
A: Experimental source: strain H37RV  
C: Gene: Rv0341  
C: Superfamily: Phasolus glycine-rich cell wall protein 1.8

Query Match 4.2%; Score 102.5; DB 2; Length 479;  
Best Local Similarity 24.8%; Pred. No. 6.1; Indels 91; Gaps 13;  
Matches 77; Conservative 30; Mismatches 115; Indels 91; Gaps 13;

QY 237 DFLTFLDRDDGAKKVFYIETREPRERLLTAHLLFVAPHNDSATG---PEASGSGP 293  
Db 6 DYILSFRSEDAARSFVAAPGR-----AMTSAGLIDIAHQISSVAANVPLNLGAGD 59  
QY 294 PSCGALGPALFASRVPGQVVVVAERDGRRLPAAVHVSITLSEEAAGAPLTA--- 350  
Db 60 PMSGL---ROAVAAHGFADQVANVGA-CD---AGAGVASVITTDVAGLAGSLGAGPL 112  
QY 351 -QGTI-----LINEVLASCYAVIEHSHWAHRAFPRL-----AHALLAAL 390  
Db 113 GGGGLAALAASSGGFGGQVGLAAQVGLGTAIVIEAVGA-QVGAGLGIGTGLGAQAGMGFG 171  
QY 391 APARTDRGGDSG-----GDRGGGGRVALTPGAADAPGAGATAGIHWTYS 436  
Db 172 GGVGLGLGGQAGGVIGGSAGAGAGAGVGGRLGNGQIGVAGQGA---GAGVGAGVGGQA 228  
QY 437 QLLYQIG-----TWLIDSE---ALHPLGNAVKSS 462  
Db 229 GIASQIGVSAGGLGGVGVNVSGLTVSSNAVLASNAQAGLIASGALNGAHPHLSS 268  
QY 463 XSRGAGGGAREGA 475  
Db 289 PLAGVGVGGQAGA 301

RESULT 46  
T27975  
hypothetical protein ZK678.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T27975  
R: Kershaw, J.  
Submitted to the EMBL Data Library, August 1996  
A: Reference number: Z20449  
A: Accession: T27975

A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-481 <WIL>  
A: Cross-references: EMBL: Z79605; PIDN: CAB01902.1; GSPDB: GN00028; CESP: ZK678.5  
A: Experimental source: clone ZK678  
C: Genetics:  
A: Gene: CESP: ZK678.5  
A: Map position: X  
A: Introns: 33/3; 99/3; 131/2; 211/1; 324/3; 342/1; 381/3; 436/3

Query Match 4.2%; Score 102.5; DB 2; Length 481;  
Best Local Similarity 19.0%; Pred. No. 6.1; Indels 105; Gaps 5;  
Matches 39; Conservative 25; Mismatches 36; Indels 105; Gaps 5;

QY 188 ENSVAAKSGS-----CFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGR 232  
Db 321 QNPVPAAPAGYAPMGAPAGSLQLIYCFGDAMVNVYNGGFKEMDELAAGDWQA----- 374  
QY 233 LLYSDFLTFLDRDDGAKKVFYIETREPRERLLTAHLLFVAPHNDSATGPEASSGSG 292  
Db 375 -----LDKNGSKSQMW-----QRTKVL-----DINIV-----QKTGIYSPMTSRG 410  
QY 293 PSCGALGPALFASRVPGQVVVVAERDGRRLPAAVHVSITLSEEAAGAPLTAQ 352  
Db 387 -----QRTKVL-----DINIV-----QKTGIYSPMTSRG 410  
QY 353 TILINRVLASCYAVIEHSHWAHRAFP 377  
Db 411 HLLVDRIRHACHSETDNYSLQNTFF 435

RESULT 47  
B75625  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: B75625  
R: White, O.; Bisen, J.A.; Heideberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamachew, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma Science 286, 1571-1577, 1999  
A: Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A: Reference number: A75250; MUID: 20036896; PMID: 10567266  
A: Accession: B75625  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-1145 <WHI>  
A: Cross-references: GB: AF001826; NID: 96450827; PIDN: AAF12641.1; PID: 96460937; TIGR: DRB00  
A: Experimental source: strain R1  
C: Genetics:  
A: Gene: DRB0065  
A: Map position: megaplasmid  
A: Genome: plasmid  
A: Note: plasmid MPl

Query Match 4.1%; Score 102; DB 2; Length 1145;  
Best Local Similarity 19.9%; Pred. No. 21;  
Matches 88; Conservative 51; Mismatches 167; Indels 136; Gaps 18;

QY 90 ENTGADRLMTORCKDLNALAISVMNQWPGVKLRVTGWDGDGHSESLHYEGRAYDIT 149  
Db 644 QSNRAHL-----RNDLQALATSMDSLESITVTPAGNPDSDA-----LTGQPYQVL 690  
QY 150 TSDRDRSKY--GMLARLAVEAGFDWVYVESKAHICSVKAENSVAAKSGGCFPGSATVHL 207  
Db 691 TDNGDSVYVGDKIAKDLVAPADQWQKFRATVIVGALPAQTCTVSNL-SVTPDNST--L 747  
QY 208 EGGTKLVKDL-----SFGDRVLAADDQGRLLYSDFLTFLDRDDGAKKVF 252  
Db 748 PVGGTQTFKAIYTTSPPCAVNVVSWSSDSPSKLTVDERG----- 786  
QY 253 YVIETREPRERLLTAHLLFVAPHNDSATGPEASSGSGPSPGALGPALFASRVPG 312

Db 787 --VARBAAGTTLTAT-----VTP-----GSGAPVSASTGVTVSAPVTSLEN 827

Qy 313 QRYVVAERDGRLLPAA-----VSVTLSEEAAGAVAPLTAQGTI--- 354

Db 828 VAVYRVG--DGSAKLSAATPVFVDIINPDGSLVRLTSLPTSGSNRFLASGTASSE 885

Qy 355 -LINRVLASCYAVIEHSHWAHRAFAFRHLAALLAALAPARTDRGG-----DSGGG 404

Db 886 GMLTRSADAHYLVITGYA-ASVGLANVKASDSATTPRVIGRLDSAGTVDTITLTPDAYSG 944

Qy 405 DR-----GGGGRRVALTPAGAADAPGAGATAGIHWSYQLLYQIGTWLLDSE 450

Db 945 DNIRSAASTDGRSFYVTGGNGSVRLYALGASSLSISSAA-----SNLRQLGIF---GD 995

Qy 451 ALHPLGVAVKSSXSRGAGGAR 472

Db 996 QLY-----VSSGSGNGTK 1008

RESULT 48

T00373

hypothetical protein KIAA0649 - human

C:Species: Homo sapiens (man)

C>Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000

C:Accession: T00373

R: Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.

DNA Res. 5, 169-176, 1998

A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete

A:Reference number: Z14142; MUID:98403880; PMID:9734811

A:Accession: T00373

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1209 <ISH>

A:Cross-references: EMBL:AB014549; NID:G3327111; PIDN:BA31624.1; PID:G3327112

A:Experimental source: brain

C:Genetics:

A:Note: KIAA0649

Query Match 4.1%; Score 102; DB 2; Length 1209;

Best Local Similarity 20.9%; Pred. No. 22;

Matches 106; Conservative 52; Mismatches 220; Indels 128; Gaps 20;

Qy 14 SSLVCSGLACRG-----RGFGKRPKLTPLAYKQFIPNVAEKLGLSGRVEGKISR 68

Db 644 SEALGGGTARGPDTMWSQGGTKDEARRLD-----EK-----ES 679

Qy 69 NSERFKELTPNYPDIIFKDENTGADRLMTQCKD-----KLNALAISVMNQWPG 119

Db 680 SEDKSSLSDEDLDTAKD--LLRSKKLKKRCREPRACRKKVRFSTAQTHFLEQLGG 737

Qy 120 VKLRVTEGWDEDDHSESLHYEGRVDDITTSDDR-----SKYGMRLARLAVEAGF 170

Db 738 LR-----RDWKDRGPVFLKSC-----LSKSKEDSGEGPKPKPPSVFGSTAERMQEG- 784

Qy 171 DWVYVESKATHCHCSKAENSVAAXSGGCFPGSATVHLEQGTKLVDLSFGDRVLAADDQ 230

Db 785 ---RASQDAALAFVRVPASASASEGNPPFPES-----QGPAPSGSLSDSSSDSDSDS 836

Qy 231 GRLLYSDFLTFLDRDDGAKKVFYVETREPRERILLTA-----AHLFVAPH 277

Db 837 IELEIRKFLA-----EKAKESVSSSEVQAEGTALGTGFPARPEVLCRKEFPAPPGVCTR 891

Qy 278 NDSATGEPEASG-SGPPSGGALGPRALFASRVPGORVYVVAERDGRRLPAAVHSV 336

Db 892 SQRAGVPHLAEGLRGTESAGAQTAGLFSQ-----GKGLPAAARGD-----PVPPRSTS 943

Qy 337 LSEBAAGAYAPLTAQGTILINRVLASCYAVIEHSHWAHRAFA-----FRLAHLAAL 390

Db 944 GGVSAGK-----LSVSRNRVYVHKDQSPRGAEFAAKSAFQGLPSC 983

Qy 391 APARTDRGSDSGGDRGGGRVALTPAGAADAPGAGATAGIHWSYQLLYQI---GWLL 447

Db 984 ATAGTEAGGAGRTFMGCGSPSFLTPSPGAERDAGAQADRTPPW-SDFAHQSRFLPSFWL 1042

Qy 448 DSEALHPLGM-AVKSSXSRGAGGAR 472

Db 1043 RSEGRDAVWRGGVGSERDKSGEPAR 1068

RESULT 49

B94339

hypothetical protein Vngi1880c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: B94339

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of Halobacterium species NRC-1

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: B94339

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-506 <STO>

A:Cross-references: CB:A5004437; NID:G10581323; PIDN:AAG20078.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG1880C

Query Match 4.1%; Score 101.5; DB 2; Length 506;

Best Local Similarity 22.6%; Pred. No. 7.8;

Matches 82; Conservative 39; Mismatches 127; Indels 115; Gaps 19;

Qy 131 DGHSEESLHYEGRVDDIT-----SDRDSKYGMRLARLAVEAGFDWVYVESKAHLHC 183

Db 25 DPAHSAVAIR-DGRIVGLSNDYDAFRADTGTREVDLGGRVVPGFVD-----AHTHL 76

Qy 184 SVKAENSVAAX-SGGCFPGSATVHLEQGTKLVDLSFGDRVLAADDQGRLL----- 234

Db 77 AVLKGHGVHADLRGADSPAAATARLAERARE-----TDAGWVLGFGYDSDQ 122

Qy 235 YSDFLTFDRDDGAKKVFYVETREP-----RRLTLTAHLFVAPHNDSATGEPEASGS 291

Db 123 WSAALSTADLD-----AVSEARPVAAIREDMTATVNGVALABHGDDEM---PAADVG- 171

Qy 292 GPPSGGALGPRALFASRVPGORVYVVAERD--GDRRLPAA-----AVHSVTL 338

Db 172 ---EDGRIVEAV-----EAVVDATDPLAGTRPLIEAAQREANERGVTAVHEMVRD 220

Qy 339 EBAAGAYAPLTAQGTILINRVLASCYAVIEHSHWAHRAFAFRHLAALLAALAPARTDRG 398

Db 221 SHAPRVYRELDAAAGELSVR-----VRLNYWAD-----HLDAVLETGLV---TNHG 262

Qy 399 GD--SGGDR-----GGGGRRVALTPAGAADAPGAGATAGIHWSYQLLYQIGTWLLDSEAL 452

Db 263 GGNVTVGKITYDGLSGRTAKLSEPYADAPG-----ETGQWVVDPSAL 307

Qy 453 HPL 455

Db 308 AAL 310

RESULT 50

C70584

probable serine-threonine protein kinase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 21-Jun-2002

C:Accession: C70584

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

Search completed: March 29, 2004, 18:34:07  
Job time : 28 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2407	97.5	462	1	SHH_HUMAN	Q15465 homo sapien
2	2068	83.8	437	1	SHH_MOUSE	Q62226 mus musculu
3	2060	83.4	437	1	SHH_RAT	Q63673 rattus norv
4	1807	73.2	425	1	SHH_CHICK	Q91035 gallus gall
5	1539.5	62.4	432	1	SHH_CYNPY	Q90385 cynops pyrr
6	1506	61.0	418	1	SHH_BRADNO	Q92008 brachydanio
7	1498	60.7	444	1	SHH_XENLA	Q92000 xenopus lae
8	1432	58.0	416	1	TWTH_BRARE	Q90419 brachydanio
9	1382.5	51.9	408	1	IHH_CHICK	Q98938 gallus gall
10	1271	51.5	411	1	IHH_HUMAN	Q14623 homo sapien
11	1262	51.1	411	1	IHH_MOUSE	P97812 mus musculu
12	1248	50.5	409	1	IHH_XENLA	Q91612 xenopus lae
13	1213	49.1	412	1	IHH_BRARE	Q98862 brachydanio
14	1183	47.9	396	1	DHH_MOUSE	Q61488 mus musculu
15	1177	47.7	396	1	DHH_HUMAN	Q91323 homo sapien
16	1079.5	43.7	396	1	DHH1_XENLA	Q91610 xenopus lae
17	1077.5	43.6	398	1	DHH2_XENLA	Q91611 xenopus lae
18	890	36.0	421	1	HH_DROME	Q52376 droscophila
19	854	34.6	422	1	HH_DROME	P56674 droscophila
20	602	24.4	121	1	SHH_RASEL	P79858 rasbora ele
21	599	24.3	131	1	SHH_CARAU	P796991 carassius a
22	599	24.3	121	1	SHH_PUNTE	P79850 puntius tet
23	598	24.2	111	1	SHH_RASHE	P79864 rasbora ket
24	598	24.2	121	1	SHH_RASPA	P79869 rasbora pav
25	597	24.2	121	1	SHH_TANAL	P79915 tanichthys
26	595	24.1	121	1	SHH_AMBCH	P79682 amblypharyn
27	594	24.1	121	1	SHH_DANAA	O13235 danio aff.
28	594	24.1	121	1	SHH_DANAE	O13234 danio aequi
29	594	24.1	121	1	SHH_DANAT	O13238 danio aff.
30	594	24.1	121	1	SHH_DANPR	O13245 danio frank
31	594	24.1	121	1	SHH_DANKE	P79709 danio kerri
32	594	24.1	121	1	SHH_DANPU	P79717 danio pulch
33	594	24.1	121	1	SHH_DANDE	O13241 devario dev



CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).

-i- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the cell, while the N-terminal peptide remains associated with the cell surface. It is also secreted in either cleaved or uncleaved form to mediate signaling to other cells (by similarity).

-i- TISSUE SPECIFICITY: Expressed in fetal intestine, liver, lung, and kidney. Not expressed in adult tissues.

-i- PM: The C-terminal domain displays an autophagy activity and a cholesterol transferase activity. Both activities result in the cleavage of the full-length protein and covalent attachment of a cholesterol moiety to the C-terminal of the newly generated N-terminal fragment (N-product). This covalent modification appears to play an essential role in restricting the spatial distribution of the protein activity to the cell surface. The N-product is the active species in both local and long-range signaling, whereas the C-product has no signaling activity (by similarity).

-i- DISEASE: Defects in SHH are the cause of holoprosencephaly 3 (HPE3) [MIM:142945]. HPE3 is a clinically variable and genetically heterogeneous malformation in which the developing forebrain fails to correctly separate into right and left hemispheres. HPE3 is associated with several distinct facies and phenotypic variability. In the most extreme cases, anophthalmia or cyclopia is evident along with a congenital absence of the mature nose. The less severe form features facial dysmorphism characterized by ocular hypertelorism, defects of the upper lip and/or nose, and absence of the olfactory nerves or corpus callosum. The majority of HPE3 cases are apparently sporadic, although clear examples of autosomal dominant inheritance have been described. Interestingly, up to 30% of obligate carriers of HPE3 gene in autosomal dominant pedigrees are clinically unaffected.

-i- SIMILARITY: Belongs to the hedgehog family.

-i- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; WWW="http://www.infobiogen.fr/services/Chromocancer/Genes/SHHID378.html".

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L38518; AAA62179.1; -;  
EMBL; AC002484; AAB67604.1; -;  
HSSP; Q62226; 1VHH.  
MEROPS; C46.002; -;  
Genew; HGNC:10848; SHH.  
MIM; 600725; -;  
MIM; 142945; -;  
GO; GO:0007500; P:mesoderm cell fate determination; TAS.  
GO; GO:0007418; P:ventral midline development; TAS.  
InterPro; IPR009045; Hedgehog/DD\_pept.  
InterPro; IPR003587; Hedgehog hint N.  
InterPro; IPR003586; Hedgehog\_hintC.  
InterPro; IPR000320; HH\_signal.  
InterPro; IPR006141; Intein S.  
InterPro; IPR001767; Pept\_C46\_hint.  
InterPro; IPR001657; Peptidase\_C46.  
Pfam; PF01085; HH\_signal; 1.  
Pfam; PF01079; Hint; 1.  
PRINTS; PR00632; SONICHHOG.  
ProDom; P003042; HH\_signal; 1.  
SMART; SM00305; HintC; 1.  
SMART; SM00306; HintN; 1.  
PROSITE; PS50817; INTEIN\_N\_TER; 1.  
Developmental protein; Autocatalytic cleavage; Hydrolase; Protease; Signal; Lipoprotein; Palmitate; Disease mutation; Holoprosencephaly.  
SIGNAL 1 23 POTENTIAL.  
CHAIN 24 462 SONIC HEDGEHOG PROTEIN.  
CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
CHAIN 198 462 SONIC HEDGEHOG PROTEIN C-PRODUCT.

FT	DOMAIN	407	411	POLY-GLY.
FT SITE	197	198		CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE	243	243		INVOLVED IN CHOLESTEROL TRANSFER (BY SIMILARITY).
FT SITE	267	267		INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
FT ACT_SITE	270	270		ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
FT LIPID	24	24		N-palmitoyl cysteine.
FT LIPID	197	197		Cholesterol glycine ester (By similarity).
FT VARIANT	31	31		G -> R (in HPE3).
FT VARIANT	88	88		/FTid=VAR_003619.
FT VARIANT	100	100		D -> V (in HPE3; familial).
FT VARIANT	115	115		/FTid=VAR_009163.
FT VARIANT	117	117		Q -> H (in HPE3; sporadic).
FT VARIANT	117	117		/FTid=VAR_009164.
FT VARIANT	117	117		N -> K (in HPE3; familial).
FT VARIANT	117	117		/FTid=VAR_009165.
FT VARIANT	117	117		W -> G (in HPE3).
FT VARIANT	117	117		/FTid=VAR_003620.
FT VARIANT	188	188		W -> R (in HPE3).
FT VARIANT	222	222		/FTid=VAR_003621.
FT VARIANT	222	222		E -> Q (in HPE3; familial).
FT VARIANT	224	224		/FTid=VAR_009166.
FT VARIANT	224	224		D -> N (in HPE3; familial).
FT VARIANT	226	226		/FTid=VAR_009167.
FT VARIANT	226	226		V -> E (in HPE3).
FT VARIANT	236	236		/FTid=VAR_009168.
FT VARIANT	236	236		A -> T (in HPE3; familial).
FT VARIANT	236	236		/FTid=VAR_009169.
FT VARIANT	236	236		S -> R (in HPE3; familial).

Query Match 97.5%; Score 2407; DB 1; Length 462;  
Best Local Similarity 100.0%; Pred. No. 2.1e-166;  
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MULLARCLLLVLSLLVCSGLACPGFGKGRHFKLTLAYKQIPNVAEKTLGASG	60
DB	1	MULLARCLLLVLSLLVCSGLACPGFGKGRHFKLTLAYKQIPNVAEKTLGASG	60
QY	61	RYEGKISNSERFKLTNNYNDIIFKDEENTGADRLMTQCKKLNALAISSVMNQPGV	120
DB	61	RYEGKISNSERFKLTNNYNDIIFKDEENTGADRLMTQCKKLNALAISSVMNQPGV	120
QY	121	KLRVTEGWDGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH	180
DB	121	KLRVTEGWDGHHSEESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYESKAH	180
QY	181	IHCYKAENSVAKSGGCPGSGATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDEL	240
DB	181	IHCYKAENSVAKSGGCPGSGATVHLQGGTKLVKDLSPGDRVLAADDQGRLLYSDEL	240
QY	241	FLDRDDGAKKVPYVITETPRERLTLTAHLFVAPHNDSATGEPEASGSGPPSGGALG	300
DB	241	FLDRDDGAKKVPYVITETPRERLTLTAHLFVAPHNDSATGEPEASGSGPPSGGALG	300
QY	301	PRALFASRVPGQVYVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQTILLINVL	360
DB	301	PRALFASRVPGQVYVAERDGRLLPAAVHSVTLSEEAAGAYAPLTAQTILLINVL	360
QY	361	ASCVAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGDRGGGGGVALTAPGA	420
DB	361	ASCVAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGDRGGGGGVALTAPGA	420
QY	421	ADAPGAGATAGTHWYSQLLYQIGTWLLSEALHPGLMAVKSS	462
DB	421	ADAPGAGATAGTHWYSQLLYQIGTWLLSEALHPGLMAVKSS	462

RESULT 2  
SHH\_MOUSE  
ID SHH\_MOUSE STANDARD; PRT; 437 AA.

AC Q62226;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH) (HHG-1).  
 GN SHH OR HHG1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=94094334; PubMed=7916661;  
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,  
 RA McMahon J.A., McMahon A.P.;  
 RA "Sonic hedgehog, a member of a family of putative signaling  
 RT molecules, is implicated in the regulation of CNS polarity.";   
 RL Cell 75:1417-1430(1993).  
 RN [2]  
 RN REVISION TO 122.  
 RC STRAIN=C57BL/6J;  
 RA McMahon A.P.;  
 RN [3]  
 RN SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
 RX MEDLINE=95236997; PubMed=7720571;  
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
 RA Zhao R., Seldin M.F., Kallion J.F., Beachy P.A.;  
 RA "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene.";   
 RL Development 120:3339-3353(1994).  
 RN [4]  
 RN FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.  
 RX MEDLINE=95254654; PubMed=7736596;  
 RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,  
 RA Beachy P.A., Jessell T.M.;  
 RA "Floor plate and motor neuron induction by different concentrations of  
 RT the amino-terminal cleavage product of sonic hedgehog  
 RL autoproteolysis.";   
 RL Cell 81:445-455(1995).  
 RN [5]  
 RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.  
 RX MEDLINE=96069744; PubMed=7477329;  
 RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;  
 RA "A potential catalytic site revealed by the 1.7-A crystal structure of  
 RT the amino-terminal signalling domain of Sonic hedgehog.";   
 RL Nature 378:212-216(1995).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESENTS THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
 CC cell, while the N-terminal peptide remains associated with the  
 CC cell surface. Is also secreted in either cleaved or uncleaved form  
 CC to mediate signaling to other cells (By similarity).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES  
 CC INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG  
 CC BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME  
 CC OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND RETINA.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.  
 CC -1- INDUCTION: By retinoic acid.  
 CC -1- PTM: The C-terminal domain displays an autoproteolysis activity  
 CC and a cholesterol transferase activity. Both activities result in

the cleavage of the full-length protein and covalent attachment of  
 a cholesterol moiety to the C-terminal of the newly generated N-  
 terminal fragment (N-product). This covalent modification appears  
 to play an essential role in restricting the spatial distribution  
 of the protein activity to the cell surface. The N-product is the  
 active species in both local and long-range signaling, whereas the  
 C-product has no signaling activity.  
 -1- SIMILARITY: Belongs to the hedgehog family.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; X76290; CAA53922.1; --  
 CC PDB; 1VHH; 29-JAN-96.  
 CC MEROPS; C46.002; --  
 CC GO; GO:0007411; P:axon guidance; IDA.  
 CC GO; GO:0001708; P:cell fate specification; IMP.  
 CC GO; GO:0007417; P:central nervous system development; IMP.  
 CC GO; GO:0001747; P:eye morphogenesis (sensu Mammalia); IMP.  
 CC GO; GO:0001654; P:eye morphogenesis; IDA.  
 CC GO; GO:0030539; P:male genital morphogenesis; IMP.  
 CC GO; GO:0007389; P:pattern specification; IMP.  
 CC GO; GO:0042127; P:regulation of cell proliferation; IDA.  
 CC InterPro; IPR009045; Hedgehog/DP rept.  
 CC InterPro; IPR003587; Hedgehog hint N.  
 CC InterPro; IPR003586; Hedgehog hint C.  
 CC InterPro; IPR000320; HH signal.  
 CC InterPro; IPR006141; Intein S.  
 CC InterPro; IPR001767; Pept\_C46\_hint.  
 CC InterPro; IPR001657; Peptidase\_C46.  
 CC Pfam; PF01085; HH signal; 1.  
 CC PRINTS; PR00632; SONICHHOG.  
 CC PRODOM; PD003042; HH signal; 1.  
 CC SMART; SM00305; HintC; 1.  
 CC SMART; SM00306; HintN; 1.  
 CC PROSITE; PS00817; INTEIN\_N\_TER; 1.  
 CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 CC Signal; Lipoprotein; Palmitate; 3D-structure.  
 CC SIGNAL; 1 24 POTENTIAL.  
 CC CHAIN 25 437 SONIC HEDGEHOG PROTEIN.  
 CC CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 CC CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 CC DOMAIN 383 387 POLY-GLY.  
 CC SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).  
 CC SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 CC SIMILARITY).  
 CC SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY  
 CC SIMILARITY).  
 CC ACT SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 CC SIMILARITY).  
 CC LIPID 25 25 N-palmitoyl cysteine (By similarity).  
 CC LIPID 198 198 Cholesterol glycine ester (By  
 CC similarity).  
 CC TURN 45 46  
 CC STRAND 48 49  
 CC TURN 54 55  
 CC TURN 57 60  
 CC TURN 69 70  
 CC HELIX 72 76  
 CC STRAND 78 79  
 CC TURN 83 84  
 CC STRAND 85 87  
 CC TURN 90 91  
 CC TURN 95 97  
 CC HELIX 98 99  
 CC STRAND 101 117  
 CC HELIX 101 117

```

PT TURN 119 120
PT STRAND 123 127
PT TURN 137 138
PT HELIX 140 143
PT TURN 144 144
PT STRAND 146 151
PT TURN 152 153
PT HELIX 156 158
PT TURN 159 168
PT STRAND 169 170
PT TURN 173 178
PT STRAND 179 180
PT TURN 181 185
PT STRAND 189 192
PT HELIX 193 193
SQ SEQUENCE 437 AA; 47773 MW; DOEB72F08E7860EF CRC64;

Query Match 83.8%; Score 2069; DB 1; Length 437;
Best Local Similarity 87.4%; Pred. No. 5.9e-142;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 60
D 1 LLLARCLLVSSLLVCSGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTILGASG 61
QY 61 RYEGKISRNSRERPKELTPNYPNDIIPKDEENTGADRLMTORCKDKLNALAISVMNQWPGV 120
D 62 RYEGKISRNSRERPKELTPNYPNDIIPKDEENTGADRLMTORCKDKLNALAISVMNQWPGV 121
QY 121 KLRVTEGWDEGHHSSESLHYEGRANDVITTSDDRSKYGMARLAVEAGFDWYYSKAH 180
D 122 KLRVTEGWDEGHHSSESLHYEGRANDVITTSDDRSKYGMARLAVEAGFDWYYSKAH 181
QY 181 IHCSVKAENSVAKSGCGFGSATVHLEQGTGLVKDLSFGDRVLAAADGRLLYSDFLT 240
D 182 IHCSVKAENSVAKSGCGFGSATVHLEQGTGLVKDLSFGDRVLAAADGRLLYSDFLT 241
QY 241 FLDRDGAQKVFYVVIETREPRERLLTAHLLFVAPNDSATGEPEASGSGSPSGALG 300
D 242 FLDRDGAQKVFYVVIETREPRERLLTAHLLFVAPNDSATGEPEASGSGSPSGALG 300
QY 301 PRALFASVRGQVYVVAERDGRRLPAVHVSITLSEEAAGAYAPLTAOGTILINRVL 360
D 287 PSALFASVRGQVYVVAERDGRRLPAVHVSITLSEEAAGAYAPLTAOGTILINRVL 346
QY 361 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDGGSDGGDRGGGGRVALTAPGA 420
D 347 ASCYAVIEHSHWAHRAFPFRLAHALLAALAPARTDGGSDGGDRGGGGRVALTAPGA 395
QY 421 ADAGAGATAGIHYSQQLYQIGTWLDSALHPLGMVAKSS 462
D 396 TEARGAETAGIHYSQQLYQIGTWLDSALHPLGMVAKSS 437

RESULT 3
SHH_RAT ID SHH_RAT STANDARD; PRT; 437 AA.
AC Q63673;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sonic hedgehog protein precursor (SHH).
GN SHH OR VHH-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:Sprague-Dawley; TISSUE:Embryonic floor plate;
RX MEDLINE=94170375; PubMed=8124714;
RA Roelink H., Augsburg A., Heenskerk J., Kozh V., Norlin S.,
RA Ruiz i Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,

```

```

PA "Floor plate and motor neuron induction by vhh-1, a vertebrate homologue
RT of hedgehog expressed by the notochord.";
RL Cell 76:761-775(1994).
CC -!- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
CC ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION
CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
CC FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
CC cell, while the N-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,
CC AND POSTERIOR LIMB BUD MESENCHYME.
CC -!- PTM: The C-terminal domain displays an autoproteolysis activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of
CC a cholesterol moiety to the C-terminal of the newly generated N-
CC terminal fragment (N-product). This covalent modification appears
CC to play an essential role in restricting the spatial distribution
CC of the protein activity to the cell surface. The N-product is the
CC active species in both local and long-range signaling, whereas the
CC C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L27340; AAA20999.1; -.
CC PIR; B53193; B53193.
CC HSP; Q62226; 1VHH.
CC MEROPS; C46.002; -.
CC InterPro; IPR009045; Hedgehog/DD_pept.
CC InterPro; IPR003587; Hedgehog_hint_N.
CC InterPro; IPR003586; Hedgehog_hint_C.
CC InterPro; IPR000320; HH_signal.
CC InterPro; IPR006141; Intein_S.
CC InterPro; IPR001767; Pept_C46_hint.
CC InterPro; IPR001657; Peptidase_C46.
CC Pfam; PF01085; HH_signal; 1.
CC Pfam; PF01079; Hint; 1.
CC PRINTS; PR00632; SONICHHOG.
CC ProDom; PD003042; HH_signal; 1.
CC SMART; SM00306; HintN; 1.
CC SMART; SM00306; HintN; 1.
CC PROSITE; PS50817; INTEIN_NTER; 1.
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
CC Signal; Lipoprotein; Palmitate.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 437 SONIC HEDGEHOG PROTEIN.
CC CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
CC CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.
CC DOMAIN 383 387 POLY-GLY.
CC SITE 198 199 CLEAVAGE (AUTO-).
CC SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
CC SIMILARITY).
CC SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
CC SIMILARITY).
CC ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY

```

Dodd J.;

"Floor plate and motor neuron induction by vhh-1, a vertebrate homologue of hedgehog expressed by the notochord.";

Cell 76:761-775(1994).

-!- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).

-!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the cell, while the N-terminal peptide remains associated with the cell surface. Is also secreted in either cleaved or uncleaved form to mediate signaling to other cells (By similarity).

-!- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE, AND POSTERIOR LIMB BUD MESENCHYME.

-!- PTM: The C-terminal domain displays an autoproteolysis activity and a cholesterol transferase activity. Both activities result in the cleavage of the full-length protein and covalent attachment of a cholesterol moiety to the C-terminal of the newly generated N-terminal fragment (N-product). This covalent modification appears to play an essential role in restricting the spatial distribution of the protein activity to the cell surface. The N-product is the active species in both local and long-range signaling, whereas the C-product has no signaling activity (By similarity).

-!- SIMILARITY: Belongs to the hedgehog family.

```

FT FT LIPID 25 25 N-palmitoyl cysteine (By similarity).
FT FT LIPID 198 198 Cholesterol glycine ester (By
FT FT LIPID 198 198 similarity).
SQ SEQUENCE 437 AA; 47630 MW; 0DBFC19FD1662A0 CRC64;

Query Match 83.4%; Score 2060; DB 1; Length 437;
Best Local Similarity 87.5%; Pred.No.2.2e-141;
Matches 405; Conservative 8; Mismatches 22; Indels 28; Gaps 4;

QY 1 MLLARCLLVVSSLLVSGLAGCGRGGRGRRPKLTPLAYKQFIPNVAEKTGLGASG 60
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2 LLLARCLFVALASSLLVCGLAGCGRGGRGRRPKLTPLAYKQFIPNVAEKTGLGASG 61
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 RYEGKISRNSRERKELTPNYPNDIIFKDBENTGADBLMTQCKDKLNALAISSVMQWPGV 120
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 RYEGKITRNSRERKELTPNYPNDIIFKDBENTGADBLMTQCKDKLNALAISSVMQWPGV 121
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KLRTVSGWEDGHHSSESLHYEGRAVDITTSORDRSKYGLARLAVEAGFDWVYYESKAH 180
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 KLRTVSGWEDGHHSSESLHYEGRAVDITTSORDRSKYGLARLAVEAGFDWVYYESKAH 181
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 IHGCVVAENSVAKSGCGPFGSATVHLEQGKLVKDLSPGRVLAADQGLLYSDFLT 240
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 IHGCVVAENSVAKSGCGPFGSATVHLEQGKLVKDLSPGRVLAADQGLLYSDFLT 241
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 FLDRDGAAGKVVYVETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 FLDRDGAAGKVVYVETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSGPPSGGALG 300
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 PRALFASRVPRGQRYVYVARDGDRLLPAHVSHVTLSEACAYAPLTAQGLILNRVL 360
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 PSLFASRVPRGQRYVYVARDGDRLLPAHVSHVTLSEACAYAPLTAQGLILNRVL 346
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 ASCVAVTEESWHAHRAFPAPRLAHALLAALAPARTDGGSGGGGGRVALTAP-G 419
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 ASCVAVTEESWHAHRAFPAPRLAHALLAALAPARTDGGSGGGGGRVALTAP-G 394
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 AADAPGAGATAGIHWSQLLYQIQTWLLDSEALHPLGMAYKSS 462
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 395 VABARGAGPAGIHWSQLLYQIQTWLLDSEALHPLGMAYKSS 437
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
SHH-CHICK STANDARD; PRT; 425 AA.
AC Q91035;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sonic hedgehog protein precursor (SHH).
GN SHH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Limb bud;
RX MEDLINE=94094333; PubMed=8269518;
RA Riddle R.D., Johnson R.L., Laufer E., Tabin C.;
RT "Sonic hedgehog mediates the polarizing activity of the ZPA.";
RL Cell 75:1401-1416(1993).
RN [2]
RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE=95254654; PubMed=7736596;
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of
RT the amino-terminal cleavage product of sonic hedgehog
RT autoproteolysis.";
RL Cell 81:445-455(1995).
```

```

-|- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
cell, while the N-terminal peptide remains associated with the
cell surface. Is also secreted in either cleaved or uncleaved form
to mediate signaling to other cells (By similarity).
-|- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD
MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE
OF THE NEURAL TUBE.
-|- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE
INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE
EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF
POLARIZING ACTIVITY (ZPA).
-|- INDUCTION: By retinoic acid.
-|- PTM: The C-terminal domain displays an autoproteolysis activity
and a cholesterol transferase activity. Both activities result in
the cleavage of the full-length protein and covalent attachment of
a cholesterol moiety to the C-terminal of the newly generated N-
terminal fragment (N-product). This covalent modification appears
to play an essential role in restricting the spatial distribution
of the protein activity to the cell surface. The N-product is the
active species in both local and long-range signaling, whereas the
C-product has no signaling activity.
-|- SIMILARITY: Belongs to the hedgehog family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; L28099; AAA72428.1; -
PIR; A49424; A49424.
HSP; O62226; 1VHH.
MEROPS; C46.001.
InterPro; IPR009045; Hedgehog/DD_pept.
InterPro; IPR003587; Hedgehog_hint_N.
InterPro; IPR003586; Hedgehog_hint_C.
InterPro; IPR00320; HH_signal.
InterPro; IPR006141; Intein_S.
InterPro; IPR001767; Pept_C46_hint.
InterPro; IPR001657; Peptidase_C46.
Pfam; PF01085; HH_signal; 1.
Pfam; PF01079; Hint; 1.
PRINTS; PR00632; SONICHHOG.
ProDom; PD003042; HH_signal; 1.
SMART; SM00305; HintC; 1.
SMART; SM00306; HintN; 1.
PROSITE; PS50817; INTEIN_N_TER; 1.
Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
Signal; Lipoprotein; Palmitate.
SIGNAL 1 26 POTENTIAL.
CHAIN 27 425 SONIC HEDGEHOG PROTEIN.
CHAIN 27 200 SONIC HEDGEHOG PROTEIN N-PRODUCT.
CHAIN 201 425 SONIC HEDGEHOG PROTEIN C-PRODUCT.
DOMAIN 390 393 POLY-THR.
SITE 200 201 CLEAVAGE (AUTO-) (PROBABLE).
SITE 246 246 INVOLVED IN CHOLESTEROL TRANSFER (BY
SIMILARITY).
SITE 270 270 INVOLVED IN AUTO-CLEAVAGE (BY
SIMILARITY).
```

```

FT ACT_SITE 273 273 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT SIMILARITY).
FT LIPID 27 27 N-palmitoyl cysteine (By similarity).
FT LIPID 200 200 Cholesterol glycine ester (By
FT similarity).
SQ SEQUENCE 425 AA; 46474 MW; DA9627443DA0173 CRC64;

Query Match 73.28; Score 1807; DB 1; Length 425;
Best Local Similarity 77.8%; Pred. No. 3.9e-123;
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 1 MLLARCLLVSVLLVCSGLACGPGRGKRRHPKLLTFLAYKQFIPNVAEKTILGASG 60
DB 4 MLLITRLLVGVICALLVSGLTGCGRGIGKRRHPKLLTFLAYKQFIPNVAEKTILGASG 63
QY 61 RYEGKISRNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
DB 64 RYEGKITRNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 123
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGFDWVYVESKAH 180
DB 124 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGFDWVYVESKAH 183
QY 181 IHCVSKAENSVAAKSGCGFPGSATVHLEQGGTKLVKDLSPGDRVLAADGGRLLYSDFLT 240
DB 184 IHCVSKAENSVAAKSGCGFPGSATVHLEHGGTKLVKDLSPGDRVLAADGGRLLYSDFLT 243
QY 241 FLDRDDGAKVYVYIETREPERLLTAALLFVAP-HNDS-ATGPEASGSGPPSGGA 298
DB 244 FLDRMDSKLFYVYIETRPARLLTAALLFVAPQHNSEATG-----STSG----- 292
QY 299 LGPRALPASVRPQORYVVAERDGRRLPAAVHVSYTLSEAGAYAPLTAQGTILINR 358
DB 293 ---QALPASNVKQORYVVLGE--GGQLLPASVHVSLSREASGAYAPLTAQGTILINR 347
QY 359 VLACYAVIEHSHWAHAFAPFRLAALLAALAPARTDRGDSGGDRGGGGRVALTAP 418
DB 348 VLACYAVIEHSHWAHAFAPFRLAALLAALAPARTDRGDSGGDRGGGGRVALTAP 418
QY 419 GAADAPGAGATAGTHWYSQQLYQITGTLWLLSEALHPLGMVAKGS 462
DB 382 DGAIPTAATTTGTHWYSRLLYRIGSVLWDGDLHPLGMVAPAS 425

RESULT 5
SHH_CYNPY
ID SHH_CYNPY STANDARD; PRT; 432 AA.
AC Q90385;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sonic hedgehog protein precursor (SHH).
GN SHH.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
OX NCBI_TaxID=8330;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=96136334; PubMed=8573168;
RA Takabatake T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.;
RT "Activation of two Cynops genes, fork head and sonic hedgehog, in
RT animal cap explants."; Commun. 218:395-401(1996).
RL Biochem. Biophys. Res. Commun. 218:395-401(1996).
CC -!- FUNCTION: INTERCELLULAR SIGNAL. ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
CC REPRESSED THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY

```

```

CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
CC cell, while the N-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC -!- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR
CC (BFGF) AND FORK HEAD.
CC -!- PTM: The C-terminal domain displays an autoproteolytic activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of
CC a cholesterol moiety to the C-terminal of the newly generated N-
CC terminal fragment (N-product). This covalent modification appears
CC to play an essential role in restricting the spatial distribution
CC of the protein activity to the cell surface. The N-product is the
CC active species in both local and long-range signaling, whereas the
CC C-product has no signaling activity (By similarity).
CC -!- SIMILARITY: Belongs to the hedgehog family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D63339; BAA09657.1; -
CC HSSP; O62226; 1VHH.
CC
CC InterPro; IPR009045; Hedgehog/DD_pept.
CC InterPro; IPR003587; Hedgehog_hint_N.
CC InterPro; IPR003586; Hedgehog_hint_C.
CC InterPro; IPR000320; HH_signal.
CC InterPro; IPR006141; Intein_S.
CC InterPro; IPR001767; Pept_C46_hint.
CC InterPro; IPR001657; PeptIdase_C46.
CC Pfam; PF01085; HH_signal; 1.
CC Pfam; PF01079; Hint; 1.
CC PRINTS; PR00632; SONICHHOG.
CC ProDom; PD003042; HH_signal; 1.
CC SMART; SM00305; HintC; 1.
CC SMART; SM00306; HintN; 1.
CC PROSITE; PS50817; INTEIN_N_TER; 1.
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
CC Signal; Lipoprotein; Palmitate.
CC SIGNAL 1 26 POTENTIAL.
CC CHAIN 27 432 SONIC HEDGEHOG PROTEIN.
CC CHAIN 27 200 SONIC HEDGEHOG PROTEIN N-PRODUCT.
CC CHAIN 201 432 SONIC HEDGEHOG PROTEIN C-PRODUCT.
CC SITE 200 201 CLEAVAGE (AUTO-) (BY SIMILARITY).
CC SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
CC SIMILARITY).
CC ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
CC SIMILARITY).
CC LIPID 27 27 N-palmitoyl cysteine (By similarity).
CC LIPID 200 200 Cholesterol glycine ester (By
CC similarity).
CC SQ SEQUENCE 432 AA; 47847 MW; B455C7E746C8E5A8 CRC64;

Query Match 62.4%; Score 1539.5; DB 1; Length 432;
Best Local Similarity 67.7%; Pred. No. 8.1e-104;
Matches 315; Conservative 36; Mismatches 75; Indels 39; Gaps 8;

QY 1 MLLARCLLVSVLLVCSGLACGPGRGKRRHPKLLTFLAYKQFIPNVAEKTILGASG 60
DB 4 MLLRRLVLAGFICALLVPSGLSCGPGRGIGTRKFKLLTFLAYKQFIPNVEKTLGASG 63
QY 61 RYEGKISRNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 120
DB 64 RYEGKITRNSERFKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAI SVMNQPGV 123
QY 121 KLRVTEGWDEGHSESLHYEGRAVDITTSDDRSKYGMRLAVEAGFDWVYVESKAH 180

```

Db 124 KURVTEGWEDDGHFEESLHYEGRVAVDITTSRDRSKYKMLARLAABAGFDWVYFESKAH 183  
QY 181 IHC5VKAENSVAAGSGCGPGSATVHEGGTKLVKD:SPGDRVLAADDQGRLLYSDFLT 240  
Db 184 IHC5VKAENSVAAGSGCGPGSATVHEGGTKLVKD:SPGDRVLAADDQGRLLYSDFLT 243  
QY 241 FLDRDDGAKVUVVITRPRRLILTAHLLFVA---PHNDSATGEPEASSGSGPPSGG 297  
Db 244 FMDKEBTVRKVFVIET--SRERVUTAHLHFVGAHEGND-----SGGDRFSVFG 293  
QY 298 ALGPRALFASRVPRGORVVVAERDGRDLPAAVHVSHTLSEAAAGAYAPLTAQOTILIN 357  
Db 294 SAGFRSMFASVRACHRVLTV-DREG-RGLREATVERVYL-EAATGAYAPVTAHGTVID 350  
QY 358 RVLASCYAVIEBSWAHRAFAPELLAHALLAALAPARTDGGDGGGGRGGRVALTA 417  
Db 351 RVLASCYAVIEBSWAHRAFAPELLAHALLAALAPARTDGGDGGGGRGGRVALTA 417  
QY 418 PGAADAFGAGATAGIHWISOLLYQIGTWLLDSEALHPLGMAYKSS 462  
Db 393 -----PPAPQSGVGHVSEILYRICTVWLQEDTHPLGMAKSS 432  
RESULT 6  
SHH\_BRARE  
ID SHH\_BRARE STANDARD; PRT; 418 AA.  
AC Q92008; O13170; O13171;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Sonic hedgehog protein precursor (SHH) (VHH-1).  
GN SHH OR VHH1.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=94170375; PubMed=8124714;  
RA Reelink H., Augburger A., Heemkerk J., Korzh V., Norlin S.,  
RA Ruiz i Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,  
RA Dodd J.; plate and motor neuron induction by vhh-1, a vertebrate homolog  
RT of hedgehog expressed by the notochord.";  
RL Curr. Biol. 7:761-775(1994).  
RN [2]  
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
RX MEDLINE=96014264; PubMed=7583153;  
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,  
RA Moon R.T., Beachy P.A.;  
RT "Patterning activities of vertebrate hedgehog proteins in the  
RL developing eye and brain.";  
RL Curr. Biol. 5:944-955(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96083328; PubMed=7579523;  
RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,  
RA McMahon A.P., Tabin C., Ingham P.W.;  
RT "The hedgehog gene family in Drosophila and vertebrate development.";  
RL Development Suppl. 43-51(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99225458; PubMed=10207136;  
RA Muller P., Chang B., Albert S., Fischer N., Tora L., Strahle U.;  
RT "Intronic enhancers control expression of zebrafish sonic hedgehog in  
RT floor plate and notochord.";  
RL Development 126:2103-2116(1999).  
RN [5]  
RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.  
RC TISSUE=Muscle;  
RX MEDLINE=97075114; PubMed=8917540;

Zardova R., Abouheif E., Meyer A.;  
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
RL closely related to the zebrafish.";  
Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE  
CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING  
CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS  
CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)  
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO  
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,  
CC PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY  
CC SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
CC cell, while the N-terminal peptide remains associated with the  
CC cell surface. Is also secreted in either cleaved or uncleaved form  
CC to mediate signaling to other cells (By similarity).  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL  
CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN  
CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT  
CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE DIENCEPHALON.  
CC -!- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF  
CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF  
CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE  
CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING  
CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE  
CC PRESOMITIVE MIDBRAIN.  
CC -!- PTM: The C-terminal domain displays an autoprotoleolysis activity  
CC and a cholesterol transferase activity. Both activities result in  
CC the cleavage of the full-length protein and covalent attachment of  
CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the  
CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC -!- SIMILARITY: Belongs to the hedgehog family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; L27585; AAA20998.1; -;  
CC EMBL; U30711; AAC59742.1; -;  
CC EMBL; Z35669; CA84738.1; -;  
CC EMBL; AF124382; AAD47913.1; -;  
CC EMBL; U51351; AAB38575.1; -;  
CC EMBL; U51370; AAB38593.1; -;  
CC FIC; A53193; A53193.  
CC HSP; Q62226; 1VHH.  
CC MEROPS; C46.002; -;  
CC ZFIN; ZDB-GENE-980526-166; shh.  
CC InterPro; IPR009045; Hedgehog/DD\_pept.  
CC InterPro; IPR003587; Hedgehog\_hint\_N.  
CC InterPro; IPR003586; Hedgehog\_hint\_C.  
CC InterPro; IPR000320; HH\_signal.  
CC InterPro; IPR006141; Intein\_S.  
CC InterPro; IPR001767; Pept\_C46\_hint.  
CC InterPro; IPR001657; Peptidase\_C46.  
CC Pfam; PF01085; HH\_signal; 1.  
CC Pfam; PF01079; Hint; 1.  
CC PRINTS; PR00632; SONICHHOG.  
CC PRODOM; PD003042; HH\_signal; 1.  
CC SMART; SM00305; HintC; 1.  
CC SMART; SM00306; HintN; 1.  
CC PROSITE; PS50817; INTEIN\_NTER; 1.  
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
KW Signal; Lipoprotein; Palmitate.  
1 23  
SIGNAL 1 POTENTIAL.

```
FT CHAIN 24 418 SONIC HEDGEHOG PROTEIN.
FT CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 198 418 SONIC HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 197 198 CLEAVAGE (AUTO-).
FT SITE 243 243 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 243 243 SIMILARITY).
FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY
FT SITE 267 267 SIMILARITY).
FT ACT_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT ACT_SITE 270 270 SIMILARITY).
FT LIPID 24 24 N-palmitoyl cysteine (By similarity).
FT LIPID 197 197 Cholesterol glycine ester (By
FT LIPID 197 197 similarity).
SQ SEQUENCE 418 AA; 46402 MW; CF000AFDFD2F5795 CRC64;
Query Match 61.0%; Score 1506; DB 1; Length 418;
Best Local Similarity 64.5%; Pred. No. 2e-101;
Matches 300; Conservative 47; Mismatches 68; Indels 50; Gaps 5;
QY 1 MLLARCLLVVSLVSLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 1 MRLTRVLLVSLVSLVCSGLACGPGRGFKRHPKLTPLAYKQFIPNVAEKTILGASG 60
QY 61 RYEGKIRNBERFKELTPNPNPDIIFKDEBNTGADRLMTORCKDKLNLAISVMQWPGV 120
DB 61 RYEGKIRNBERFKELTPNPNPDIIFKDEBNTGADRLMTORCKDKLNLAISVMQWPGV 120
QY 121 KLRVTEGWDEGHSEFSLHYEGRVADITTSDDRSKYGMRLAVEAGFDWYVESKAH 180
DB 121 KLRVTEGWDEGHSEFSLHYEGRVADITTSDDRSKYGMRLAVEAGFDWYVESKAH 180
QY 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGRVLAADGGRLLYSDFLT 240
DB 181 IHCSVKAENSVAKSGCGFPGSATVHLEQGGTKLVKDLSPGRVLAADGGRLLYSDFLT 240
QY 241 FLDDDDGAKKVFYIETEPERLLTAHLLFVAPENDSATGEPEASSGGPPSGGALG 300
DB 241 FLDDDDGAKKVFYIETEPERLLTAHLLFVAPENDSATGEPEASSGGPPSGGALG 300
QY 241 FTDRSDTTRRVFYIETQEPVEKILTAAHLLFVLT-----DNSTDLHTMT----- 286
DB 241 FTDRSDTTRRVFYIETQEPVEKILTAAHLLFVLT-----DNSTDLHTMT----- 286
QY 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSSEAGAYAPLTAQGTILINRVL 360
DB 301 PRALFASRVPRGQVYVVAERDGRLLPAAVHSVTLSSEAGAYAPLTAQGTILINRVL 360
QY 287 --AAVASSVRAGQKVVV---DSDGQKSLVIVQIVTESQSGSFAPVTAHGTIVVDRL 340
DB 287 --AAVASSVRAGQKVVV---DSDGQKSLVIVQIVTESQSGSFAPVTAHGTIVVDRL 340
QY 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDRGGSDGGDRGGGGRRVALTAPGA 420
DB 361 ASCVAVIEHSHWAHRAFPRLAHALLAALAPARTDRGGSDGGDRGGGGRRVALTAPGA 420
QY 341 ASCVAVIEDQGLAHAFAPARLYTVVSFFPPQNS----- 375
DB 341 ASCVAVIEDQGLAHAFAPARLYTVVSFFPPQNS----- 375
QY 421 ADAPGAGAT---AGHWYSOLLYQIGTWLDDSEALHPLGMVAKSS 462
DB 376 --SRSNATLQOEGVHWYSRLLYQMGITWLLDSNMLHPLGMSVNS 418
RESULT 7
SHH_XENLA STANDARD; PRT; 444 AA.
AC Q92000; Q91894;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sonic hedgehog protein precursor (X-SHH) (VHH-1).
GN SHH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95357169; PubMed=7630736;
RA Stoll M.A., Shi Y.-B.;
RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis
and thyroid hormone-dependent metamorphosis.";
```

15-JUL-1999 (Rel. 38, Created)  
13-JUL-1999 (Rel. 38, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
TIGgy-winkle hedgehog protein precursor (TWHH).  
GN TWHH.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
[1]  
RN SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
RP TISSUE=Embryo;  
RC MEDLINE=96014264; PubMed=7583153;  
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,  
RA Porter J.A., Moon R.T., Beachy P.A.;  
RT "Patterning activities of vertebrate hedgehog proteins in the  
RT developing eye and brain.";  
RL Curr. Biol. 5:944-955(1995).  
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL  
CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING  
CC EYES.  
CC -1- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
CC cell, while the N-terminal peptide remains associated with the  
CC cell surface. Is also secreted in either cleaved or uncleaved form  
CC to mediate signaling to other cells (By similarity).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL  
CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN  
CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE  
CC DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.  
CC -1- PM: The C-terminal domain displays an autoprolysis activity  
CC and a cholesterol transferase activity. Both activities result in  
CC the cleavage of the full-length protein and covalent attachment of  
CC a cholesterol moiety to the C-terminal of the newly generated N-  
CC terminal fragment (N-product). This covalent modification appears  
CC to play an essential role in restricting the spatial distribution  
CC of the protein activity to the cell surface. The N-product is the  
CC active species in both local and long-range signaling, whereas the  
CC C-product has no signaling activity (By similarity).  
CC -1- SIMILARITY: Belongs to the hedgehog family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: U30710; AAC59741.1; --  
CC HSSP: Q62226; LVHH.  
CC MESOPS; C46.001; --  
CC ZFIN; ZDB-GENE-980526-41; twhh.  
CC InterPro: IPR009045; Hedgehog/DD\_pept.  
CC InterPro: IPR003587; Hedgehog\_hint\_N.  
CC InterPro: IPR003586; Hedgehog\_hint\_C.  
CC InterPro: IPR000320; HH\_signal.  
CC InterPro: IPR001767; Rept\_C46\_hint.  
CC InterPro: IPR001657; Peptidase\_C46.  
CC Pfam: PF01085; HH\_signal; 1.  
CC Pfam: PF01079; Hint; 1.  
CC PRINTS; PR00632; SONICHHOG.  
CC ProDom: PD003042; HH\_signal; 1.  
CC SMART; SM00305; HintC; 1.  
CC SMART; SM00306; HintN; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
CC Signal; lipoprotein; Palmitate.  
CC SIGNAL 1 26 POTENTIAL.  
FT CHAIN 27 416 TIGGY-WINKLE HEDGEHOG PROTEIN.  
FT CHAIN 27 416 TIGGY-WINKLE HEDGEHOG PROTEIN N-PRODUCT.  
FT CHAIN 201 416 TIGGY-WINKLE HEDGEHOG PROTEIN C-PRODUCT.  
FT SITE 200 201 CLEAVAGE (AUTO-).  
FT



Db 124 RVTEGDEGHHSESLHVEGRAVDITTSDDRNKYMGLARLAVEAGFDWVYVESKAH1H 183  
Qy 183 CSYKAENSAKSGGCGPFSATVHLQGGTKLVKDLSPGDRVLAAADDQGRLLYSDFLTF 242  
Db 184 CSVKSEHSAKSGGCGPFSATVHLQGGTKLVKDLSPGDRVLAAADDQGRLLYSDFLTF 243  
Qy 243 DRDGAQKVFVYETREPRERLITTAHLLFVAPHNDSATGSPPEASSGGSPGGALGPR 302  
Db 244 DKEPRALTAFHVIETROPFRRLALTTHLLFVA---DNA-----SAPAAQPR 287  
Qy 303 ALFASRVRCQRYVVAERDGRRLLPAAVHVSITLSEEAAGAYAPLTAQGTILINRVLAS 362  
Db 288 PTASHVQPGHFVLVAV---GSGGLQPAEVGVYR-GRTDVGAYAPLTRHGLTVDDVVAS 343  
Qy 363 CYAVIEBHWAPRAFPAPFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTAPGAAD 422  
Db 344 CFALVREQQLAQAFWPLRYLHSL-----GGPGVQGD----- 376  
Qy 423 APCAGATAGHMYSQLLYQIGTWTLLDSEALHPLG 456  
Db 377 -----GVHWYSGLLYLRGLRMLLPDPDFHPLG 402

## RESULT 10

IHH\_HUMAN STANDARD; PRT; 411 AA.  
ID Q14322; Q8N4B9;  
AC 15-JUL-1999 (Rel. 38, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update).  
DE 10-OCT-2003 (Rel. 42, Last annotation update).  
DE Indian hedgehog protein precursor (IHH) (HHG-2).  
GN IHH.  
OS Homo sapiens (Human).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN RP  
RA Tate G., Kishimoto K., Mitsuwa T.;  
RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
human cancer cell lines and embryonic organs";  
RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
[2]  
RN RP  
RA "SEQUENCE OF 95-411 FROM N.A."  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Abbey J., Helton E., Kettelman M., Vadan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
RN RP  
RA "SEQUENCE OF 100-411 FROM N.A."  
RX TISSUE=Fetal lung;  
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,  
RA Gastier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,  
RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;

"Cloning, expression, and chromosomal location of SHH and IHH: two  
human homologues of the Drosophila segment polarity gene hedgehog.";  
RL Genomics 28:44-51(1995).  
[4]  
RN RP  
RA "SEQUENCE OF 124-172 FROM N.A."  
RX MEDLINE=95236997; PubMed=7720571;  
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;  
RT "Products, genetic linkage and limb patterning activity of a murine  
hedgehog gene.";  
RL Development 120:3339-3353(1994).  
[5]  
RN RP  
RA "VARIANTS BDA1 LYS-95; GLU-100 AND LYS-131."  
RX MEDLINE=21372448; PubMed=11455389;  
RA Gao B., Guo J., She C., Shu A., Yang M., Tan Z., Yang X., Guo S.,  
RA Feng G., He L.;  
RT "Mutations in IHH, encoding Indian hedgehog, cause brachydactyly type  
A-1.";  
RL Nat. Genet. 28:386-388(2001).  
[6]  
RN RP  
RA "VARIANT BDA1 ASN-100."  
RX MEDLINE=22271632; PubMed=12384778;  
RA McCready M.E., Sweeney E., Fyfe A.E., Donnai D., Baig A., Racacho L.,  
RA Warman M.L., Hunter A.G.W., Sulman D.E.;  
RT "A novel mutation in the IHH gene causes brachydactyly type A1: a  
95-year-old mystery resolved.";  
RL Hum. Genet. 111:368-375(2002).  
[7]  
RN RP  
RA "VARIANTS ACFD LEU-46 AND ALA-190."  
RX MEDLINE=22534978; PubMed=12632327;  
RA Hellenans J., Coucke P.J., Giedion A., De Paape A., Kramer P.,  
RA Beemer F., Mortier G.R.;  
RT "Homozygous mutations in IHH cause acrocapitofemoral dysplasia, an  
autosomal recessive disorder with cone-shaped epiphyses in hands and  
hips.";  
RL Am. J. Hum. Genet. 72:1040-1046(2003).  
CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)  
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO  
ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN  
ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH  
AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION  
OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
cell, while the N-terminal peptide remains associated with the  
cell surface. Is also secreted in either cleaved or uncleaved form  
to mediate signaling to other cells (By similarity).  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT  
KIDNEY AND LIVER.  
CC -!- PTM: The C-terminal domain displays an autotranspeptidase activity  
and a cholesterol transferase activity. Both activities result in  
the cleavage of the full-length protein and covalent attachment of  
a cholesterol moiety to the C-terminal of the newly generated N-  
terminal fragment (N-product). This covalent modification appears  
to play an essential role in restricting the spatial distribution  
of the protein activity to the cell surface. The N-product is the  
active species in both local and long-range signaling, whereas the  
C-product has no signaling activity (By similarity).  
CC -!- DISEASE: Defects in IHH are the cause of brachydactyly type A1  
(BDA1) [MIM:112500]. BDA1 is an autosomal dominant disorder  
characterized by middle phalanges of all the digits rudimentary or  
fused with the terminal phalanges. The proximal phalanges of the  
thumbs and big toes are short.  
CC -!- DISEASE: Defects in IHH are a cause of acrocapitofemoral  
dysplasia (ACFD) [MIM:607778]. ACFD is a disorder characterized by  
short stature of variable severity with postnatal onset. The most  
constant radiographic abnormalities are observed in the tubular  
bones of the hands and in the proximal part of the femur. Cone-  
shaped epiphyses or a similar epiphyseal configuration with  
premature epimetaphyseal fusion result in shortening of the  
skeletal components involved. Cone-shaped epiphyses were also  
present to a variable extent at the shoulders, knees, and ankles.  
CC -!- SIMILARITY: Belongs to the hedgehog family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; AB018076; BAA33523.2; -  
 DR EMBL; AB018092; BAA33523.2; JOINED.  
 DR EMBL; AB018075; BAA33523.2; JOINED.  
 DR EMBL; BC034757; AAB34757.1; ALT\_INIT.  
 DR EMBL; L38517; AAA62178.1; -  
 DR HSSP; O62226; 1VHH.  
 DR MEROPS; C46.003; -  
 DR Genew; HGNC:5956; IHH.  
 DR MIN; 600726; -  
 DR MIN; 112500; -  
 DR MIN; 607778; -  
 DR GO; GO:0005576; C:extracellular; NAS.  
 DR GO; GO:0015485; F:cholesterol binding; NAS.  
 DR GO; GO:0005113; F:patched binding; NAS.  
 DR GO; GO:0007267; P:cell-cell signaling; NAS.  
 DR InterPro; IPR003587; Hedgehog\_hint\_N.  
 DR InterPro; IPR003586; Hedgehog\_hint\_C.  
 DR InterPro; IPR003020; HH\_signal.  
 DR InterPro; IPR006141; Intein\_S.  
 DR InterPro; IPR001677; Pept\_C46\_hint.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR Pfam; PF01079; Hint; 1.  
 DR PRINTS; PD00632; SONICHHOG.  
 DR ProDom; PD003042; HH\_signal; 1.  
 DR SMART; SMO0305; HintC; 1.  
 DR SMART; SMO0306; HintN; 1.  
 DR PROSITE; PS00817; INTEN\_N\_TER; 1.  
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate; Disease mutation.  
 FT SIGNAL 1 POTENTIAL.  
 FT CHAIN 28 411 INDIAN HEDGEHOG PROTEIN.  
 FT CHAIN 28 202 INDIAN HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 203 411 INDIAN HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 202 203 CLEAVAGE (AUTO-) (BY SIMILARITY).  
 FT SITE 248 248 INVOLVED IN CHOLESTEROL TRANSFER (BY SIMILARITY).  
 FT SITE 272 272 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).  
 FT ACT\_SITE 275 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).  
 FT LIPID 28 28 N-palmitoyl cysteine (By similarity).  
 FT LIPID 202 202 Cholesterol glycine ester (By similarity).  
 FT VARIANT 46 46 P -> L (in ACDF).  
 FT VARIANT 95 95 E -> K (in BDAL).  
 FT VARIANT 100 100 D -> E (in BDAL).  
 FT VARIANT 100 100 D -> N (in BDAL).  
 FT VARIANT 131 131 E -> K (in BDAL).  
 FT VARIANT 190 190 V -> A (in ACDF).  
 FT VARIANT 100 100 D -> R (in REF. 3).  
 FT VARIANT 246 246 F -> L (in REF. 1).  
 FT VARIANT 309 309 V -> A (in REF. 1).  
 SQ SEQUENCE 411 AA; 45250 MW; 4DA90C83F5ABF758 CRC64;  
 Query Match 51.5%; Score 1271; DB 1; Length 411;  
 Best Local Similarity 57.7%; Pred. No. 1.8e-84;  
 Matches 267; Conservative 45; Mismatches 85; Indels 66; Gaps 10;

QY 7 CILLVLVSSLLVCSGLACGPGFG-KRRHPKLTPLAYKOFIPNVAEKTLCASGRVEGK 65  
 DB 13 CLVLVLL--LVPAWAGCGPRVGVSRPRKLVPLAYKQFSPNPEKTLGASGRVEGK 70  
 QY 66 ISRNSERFKELTPNNPDIIIFKBENTGADRLMTQRCCKLNALAISSVMNMPGKLVRT 125  
 DB 71 IARSSERFKELTPNNPDIIIFKBENTGADRLMTQRCCKDLNLSLAISVMNMPGKLVRT 130  
 QY 126 EGWDEDDGHHSSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYYSKAIHCSV 185  
 DB 131 EGWDEDDGHHSSESLHYEGRAVDITTSDDRNKYGMLARLAVEAGFDWVYYSKAIHCSV 190  
 QY 186 KAENSVAAKSGCGPGSATVHLEOGTKLVKDLSPGDRVLAAADQGRLLYSDFITFLDRD 245  
 DB 191 KSEHSAAKTGCGPAGQVRLSGARVALSAVRFGDRVLAMGEDSGSPFSDVLIFLDRE 250  
 QY 246 DGAKKVFYVETREPRERELLTAHLLFVAPHNSATGEPEASSGSPSGGALGPRLAF 305  
 DB 251 PHRLRAFQVETQDPRLALTPARLLFTA---DNHT-EPAARF-----RATF 294  
 QY 306 ASRVPRGQVVYVAERDGRRLPAAVHSHVTLSEEAAGAYAPLTAQGTILINRVLASCYA 365  
 DB 295 ASHVQPGQYVLA---GVPGIQQPARVAAS-THVALGAYAPLTGHGLTVVEDVVASCF 349  
 QY 366 VIEHSHWAHRAFPAPRLAHALLAALAPARTDGGSDGGGGRGGRVALTPAGNADAPG 425  
 DB 350 AVADHHLAQLAFWPLRLFHSL-----ANGSWTPG 378  
 QY 426 AGATAGIHWYSLYQIQTWLLDSEALHPLGMVAKSSXSXRGAG 468  
 DB 379 ----EGVHWYQLLYRLGRLLJEGSFPLGMS-----GAG 410  
 RESULT 11  
 IHH\_MOUSE  
 ID IHH\_MOUSE STANDARD; PRT; 411 AA.  
 AC P97812; O61724;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (HHG-2).  
 GN IHH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
 RC STRAIN=CD-1; TISSUE=Kidney;  
 RX MEDLINE=97236802; PubMed=9079674;  
 RA Valentini R.P., Brookhiser W.T., Park J., Yang T., Briggs J.,  
 RA Dressler G., Holzman L.B.;  
 RT "Post-translational processing and renal expression of mouse indian hedgehog";  
 RL J. Biol. Chem. 272:8466-8473(1997).  
 RN [2]  
 RP SEQUENCE OF 76-411 FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=94094334; PubMed=7916661;  
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,  
 RA McMahon J.A., McMahon A.P.;  
 RT "Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity";  
 RL Cell 75:1417-1430(1993).  
 RN [3]  
 RP REVISIONS.  
 RC STRAIN=C57BL/6J;  
 RA St Jacques B.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 124-172 FROM N.A.  
 RX MEDLINE=95236997; PubMed=7720571;

RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;  
 RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene."  
 RL Development 120:3339-3353 (1994).  
 CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)  
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO  
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN  
 CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH  
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION  
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).  
 CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
 CC cell, while the N-terminal peptide remains associated with the  
 CC cell surface. Is also secreted in either cleaved or uncleaved form  
 CC to mediate signaling to other cells (By similarity).  
 CC -!- TISSUE SPECIFICITY: IN THE EMBRYO. DETECTED IN THE DEVELOPING GUT,  
 CC THE GROWTH ZONE OF CARTILAGE OF DEVELOPING LONG BONES, EPITHELIUM  
 CC AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL  
 CC CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.  
 CC -!- DEVELOPMENTAL STAGE: Detected at 10 dpc in developing gut, at 14.5  
 CC days dpc in the cartilage primordium and in the developing  
 CC urogenital sinus. Expression increases with gestational age in  
 CC kidney and duodenum, becoming maximal in adulthood.  
 CC -!- PTM: The C-terminal domain displays an autoproteolysis activity  
 CC and a cholesterol transferase activity. Both activities result in  
 CC the cleavage of the full-length protein and covalent attachment of  
 CC a cholesterol moiety to the C-terminal of the newly generated N-  
 CC terminal fragment (N-product). This covalent modification appears  
 CC to play an essential role in restricting the spatial distribution  
 CC of the protein activity to the cell surface. The N-product is the  
 CC active species in both local and long-range signaling, whereas the  
 CC C-product has no signaling activity (By similarity).  
 CC -!- SIMILARITY: Belongs to the hedgehog family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; U85610; BAB49692.1; ALT\_INIT.  
 DR EMBL; X78291; CAA53923.1;  
 DR HSP; Q62226; LVH.  
 DR MEROPS; C46.003; -.  
 DR MGD; MG1.96533; Ihh.  
 DR InterPro; IPR009045; Hedgehog/DD\_pept.  
 DR InterPro; IPR003587; Hedgehog\_hint\_N.  
 DR InterPro; IPR003586; Hedgehog\_hint\_C.  
 DR InterPro; IPR00320; HH\_signal.  
 DR InterPro; IPR006141; Intein\_S.  
 DR InterPro; IPR001767; Pept\_C46\_hint.  
 DR InterPro; IPR001657; Peptidase\_C46.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR Pfam; PF01079; Hint; 1.  
 DR PRINTS; PR006302; SONICHOG.  
 DR ProDom; PD003042; HH\_signal; 1.  
 DR SMART; SM00305; HintC; 1.  
 DR SMART; SM00306; HintN; 1.  
 DR PROSITE; PS50817; INTEIN\_N\_TPR; 1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 411 INDIAN HEDGEHOG PROTEIN.  
 FT CHAIN 28 202 INDIAN HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 203 411 INDIAN HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 202 203 CLEAVAGE (AUTO-).  
 FT SITE 248 248 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SIMILARITY).  
 FT SITE 272 272 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT SIMILARITY).

FT ACT\_SITE 275 275 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT SIMILARITY).  
 FT LIPID 28 28 N-palmitoyl cysteine (By similarity).  
 FT LIPID 202 202 Cholesterol glycine ester (By  
 FT similarity).  
 FT CONFLICT 383 383 W -> S (IN REF. 2).  
 SQ SEQUENCE 411 AA; 45485 MW; 08BE7AD8507CD9B CRC64;  
 Query Match 51.1%; Score 1262; DB 1; Length 411;  
 Best Local Similarity 56.9%; Pred. No. 8; le-84;  
 Matches 264; Conservative 46; Mismatches 86; Indels 68; Gaps 10;  
 QY 7 CLLLVSVLLVCSGLACGPRGFG-KRRHPKLLTPLAYKQFIPNVAETLHGSGRYEGK 65  
 DB 13 CLFLLLL--LLVPAARGCGPRVVGRRRPPRKLPLVLAAYKQFSPNPEKTLGASGYEGK 70  
 QY 66 ISNSERFKELTNNYPDIIFKDEENTGADRLMTQCKKLNALAI SVMNQFQVLRVT 125  
 DB 71 IASERFKELTNNYPDIIFKDEENTGADRLMTQCKKLNALAI SVMNQFQVLRVT 130  
 QY 126 EGWDEDEGHSESLHYEGRAVDITTSDRDRSKYMLARLAVEAGFDWVYVESKAHIHCSV 185  
 DB 131 EGWDEDEGHSESLHYEGRAVDITTSDRDRSKYMLARLAVEAGFDWVYVESKAHVCSV 190  
 QY 186 KAENSVAKSGCGPGSATVHLEGGTKLVKDLSPGDRVLAADDQGRLLVSDFTLDRD 245  
 DB 191 KSEHSAARKTGGCFPAGAQVRLNGERVALSVAKPGDRVLAMGEDGTPTFSDVLIFLDRE 250  
 QY 246 DGAKKVFYIETREPRERLLTAAHLFLVAPHNDSATGEPEASSGSGPPSGGALGPRALF 305  
 DB 251 PNRLARQVIETQDPPRRALTTPAHLFLIA---DNHT-EPAAHF-----RATF 294  
 QY 306 ASRVPRQRYVVAERDGRRLIPAAVSHVTLSEEAAGAVAPLTAOCTILINVLASCYA 365  
 DB 295 ASHVQPGQYVLYV---SGVPLQPARVAVS--THVALGSIAPUTRRGLTVEDVVASCPA 349  
 QY 366 VIEHSHWAHRAFAFRLAHALL-AALAPARTDRGGSGGGDRGGGRVALTPAGAADAP 424  
 DB 350 AVADHLLAQAFWPLRFPFLAWSWTPE----- 379  
 QY 425 GAGATAGIHWSYQLLYIGTWTLLDSEALHPLGNVAKSSXSRGAG 468  
 DB 380 -----GVHWYPMQLYRLGLLLEESTFHLGMS-----GAG 410  
 RESULT 12  
 ID IHH\_XENLA STANDARD; PRT; 409 AA.  
 AC Q91612;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Banded hedgehog protein) (X-  
 DE BHH)  
 GN IHH OR BHH.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=95401852; PubMed=7671800;  
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,  
 RA Moon R.T., Beachy P.A.;  
 RT "Distinct expression and shared activities of members of the hedgehog  
 RT gene family of Xenopus laevis."  
 RL Development 121:2337-2347 (1995).  
 CC -!- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF  
 CC ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC  
 CC CEMENT GLAND FORMATION IN EMBRYOS.  
 CC -!- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE

CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED  
 CC OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY  
 CC SIMILARITY).

CC -!- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA. AT  
 CC STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM. EXPRESSED  
 CC THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS  
 CC SYSTEM AND IN THE DERMATOME OF SOMITES.

CC -!- INDUCTION: BY activin.

CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
 CC and a cholesterol transferase activity. Both activities result in  
 CC the cleavage of the full-length protein and covalent attachment of  
 CC a cholesterol moiety to the C-terminal of the newly generated N-  
 CC terminal fragment (N-product). This covalent modification appears  
 CC to play an essential role in restricting the spatial distribution  
 CC of the protein activity to the cell surface. The N-product is the  
 CC active species in both local and long-range signaling, whereas the  
 CC C-product has no signaling activity (By similarity).

CC -!- SIMILARITY: Belongs to the hedgehog family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; U26404; AA85165.1; -;  
 CC HSSP; Q62226; 1VHH.  
 CC MEROPS; C46.UPW.  
 CC InterPro; IPR009045; Hedgehog/DD\_pept.  
 CC InterPro; IPR003587; Hedgehog hint N.  
 CC InterPro; IPR003586; Hedgehog hint C.  
 CC InterPro; IPR000320; HH signal.  
 CC InterPro; IPR001767; Pept C46 hint.  
 CC InterPro; IPR001657; Peptidase\_C46.  
 CC Pfam; PF01085; HH signal; 1.  
 CC Pfam; PF01079; Hint; 1.  
 CC PRINTS; PR00632; SONICHHOG.  
 CC ProDom; PDO03042; HH signal; 1.  
 CC SMART; SM00305; HintC; 1.  
 CC SMART; SM00306; HintN; 1.  
 CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 23  
 FT CHAIN 24 409  
 FT CHAIN 24 197  
 FT CHAIN 198 409  
 FT CHAIN 32 35  
 FT DOMAIN 197 198  
 FT SITE 267  
 FT SITE 267  
 FT ACT\_SITE 270 270  
 FT LIPID 24 24  
 FT LIPID 197 197  
 FT SEQUENCE 409 AA; 45591 MW; 6FB265367FB98627 CRC64;  
 Query Match 50.5%; Score 1248; DB 1; Length 409;  
 Best Local Similarity 56.3%; Pred. No. 8.3e-83;  
 Matches 256; Conservative 53; Mismatches 92; Indels 54; Gaps 7;  
 QY 4 LARCLLVLSLVCSGL-ACGPGRGKRRHPKCLTPLYAKGFIQVNAEKLIGASGRY 62  
 DB 3 LPKVVLLCAALLLSGAVRGCGRGVRRRTKLSPLSKQPSNVPEKTLIGASGRY 62  
 QY 63 EGIISRSERFKELTPNNPDIIIFKDEINTGADRLMTQCKDKLNALAIISVMNQWPGVKL 122  
 DB 63 EGIISRSERFKELTPNNPDIIIFKDEINTGADRLMTQCKDKLNALAIISVMNQWPGVKL 122  
 QY 123 RVTEGWDEDDGHSESLHYEGRAVDITTSDRDRSKYGMARLAVEAGFDWVYSEKAIH 182

DB 123 RVTEGWDEDDGHSESLHYEGRAVDITTSDRDRSKYGMARLAVEAGFDWVYSEKAIH 182  
 QY 183 CSVKAEISVAKSGCGCFPGSATVHLEQGGTKLVKDLSPCDRVLAADDOGRLLYSDFLTEL 242  
 DB 183 CSVKSEHSAATKCGCFEGEALATLESKEKIPVSQLSPGLRVLAWDNGSRPTYSDFLSFL 242  
 QY 243 DRDDCAKVFVYVETREPRERRLILLTAHLLFVAHPNDSATGEPEASSGSGPPSGALGPR 302  
 DB 243 DHSPEEHMFQVKTQDPHRLFLTPAHLIFVSDNYSFASEFQ----- 286  
 QY 303 ALFASRVPRGQVRYVVAERDGRLLPAVHSVTLSSEAAGAYAPLTAGQTLININVLAS 362  
 DB 287 AVFASVSRPGQYI-LVSNVWG---LIPAKVRSVN-TQTYGAYAPLTQHGTLWDDVVVS 341  
 QY 363 CYAVTEHSWAHRAFAFRLAHALLAALAPARTDRGSDGGGGRGGRVALTPAGAAD 422  
 DB 342 CFALVQKQLAQIVWPLRVLYNL-----GIIAGTOP----- 373  
 QY 423 AFGATAGIHWSOLLYOIGTWLILDSALHPLGM 457  
 DB 374 ----SQQMGHWYSKALYHLGLRLILHGNFHLGI 404

RESULT 13  
 ID IHH BRARE STANDARD; PRT; 412 AA.  
 AC Q98862;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Echidna hedgehog protein).  
 GN IHH OR EHH.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 CC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96310864; PubMed=8684485;  
 RA Currie P.D., Ingham P.W.;  
 RT "Induction of a specific muscle cell type by a hedgehog-like protein  
 RT in zebrafish.";  
 RL Nature 382:452-455(1996).  
 RN [2]  
 RP SEQUENCE OF 113-170 FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=97075114; PubMed=8917540;  
 RA Zardoya R., Abouheif E., Meyer A.;  
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
 RT closely related to the zebrafish.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
 CC -!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE  
 CC NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER  
 CC DIFFERENTIATION.  
 CC -!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the  
 CC cell, while the N-terminal peptide remains associated with the  
 CC cell surface. Is also secreted in either cleaved or uncleaved form  
 CC to mediate signaling to other cells (By similarity).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.  
 CC -!- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.  
 CC DISAPPEARS AT THE END OF THE SOMITOGENESIS.  
 CC -!- PTM: The C-terminal domain displays an autophosphorylation activity  
 CC and a cholesterol transferase activity. Both activities result in  
 CC the cleavage of the full-length protein and covalent attachment of  
 CC a cholesterol moiety to the C-terminal of the newly generated N-  
 CC terminal fragment (N-product). This covalent modification appears  
 CC to play an essential role in restricting the spatial distribution  
 CC of the protein activity to the cell surface. The N-product is the  
 CC active species in both local and long-range signaling, whereas the  
 CC C-product has no signaling activity (By similarity).

Dd	309	--GKLKRSQITHVGVRED-QGLYPLPTAHGTVVVNDVLITSCYAANVRQLAHMAFAPL	365
Qy	381	RLAHALLAALAFARTDRGGDSGGDGGGRVALTAPGAADAPGATAGIHWISOLLY	440
Dd	366	RLYSW-----TGPDQVLKNGLHWSOVLI	390
Qy	441	QIGTWLLDSEALHPLGM	457
Dd	391	GICLKLLDSELFPFLAL	407

RESULT 14

ID	DHH_MOUSE	STANDARD;	PRT;	396 AA.
AC	Q61488;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DN	Desert hedgehog protein precursor (DHH) (HHG-3).			
GN	DHH.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	NCBI_TaxID=10090;			
CX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/Sv;			
EX	MEDLINE=34034334; PubMed=7916661;			
RA	Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J., McMahon J.A., McMahon A.P., McManis M.F., Beachy P.A.;			
FA	"Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity.";			
RT	Cell 75:1417-1430(1993).			
RL	[2]			
RP	SEQUENCE OF 120-168 FROM N.A.			
RA	MEDLINE=95236397; PubMed=7720571;			
EX	Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K., Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;			
FA	"Products, genetic linkage and limb patterning activity of a murine hedgehog gene.";			
RT	Development 120:3339-3353(1994).			
RL	[1]			
CC	-!- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A SPERMATOCTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES DEVELOPMENT.			
CC	-!- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the cell, while the N-terminal peptide remains associated with the cell surface. Is also secreted in either cleaved or uncleaved form to mediate signaling to other cells (By similarity).			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN LIMB BUDS.			
CC	-!- PMF: The C-terminal domain displays an autoprolysis activity and a cholesterol transferase activity. Both activities result in the cleavage of the full-length protein and covalent attachment of a cholesterol moiety to the C-terminal of the newly generated N-terminal fragment (N-product). This covalent modification appears to play an essential role in restricting the spatial distribution of the protein activity to the cell surface. The N-product is the active species in both local and long-range signaling, whereas the C-product has no signaling activity (By similarity).			
CC	-!- SIMILARITY: Belongs to the hedgehog family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce or send an email to licenses@isb-sib.ch).			
CC	EMBL; X76292; CRA55924.1; -			
DR	PIR; B49425; B49425.			
DR	HSPG; Q62226; 1VEH.			

```

DR MEROPS; C46.004; -.
DR MGD; MGI:94891; Dhh.
DR InterPro; IPR003045; Hedgehog/DD_pept.
DR InterPro; IPR003587; Hedgehog hint N.
DR InterPro; IPR003586; Hedgehog hint C.
DR InterPro; IPR003020; HH signal.
DR InterPro; IPR006141; Intein S.
DR InterPro; IPR001767; Pept C46 hint.
DR InterPro; IPR001657; Peptidase_C46.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PD00632; SONICHOG.
DR ProDom; PD003042; HH_signal; 1.
DR SMART; SM00305; HintC; 1.
DR SMART; SM00306; HintN; 1.
DR PROSITE; PS00817; INTEIN_N_TER; 1.
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 22
FT CHAIN 23 396
FT CHAIN 23 198
FT CHAIN 199 396
FT SITE 198 199
FT SITE 244 244
FT SITE 268 268
FT ACT_SITE 271 271
FT LIPID 23 23
FT LIPID 198 198
FT SEQUENCE 396 AA; 43542 MW; AFPE8051BE950FDB CRC64;
Query March 47.9%; Score 1183; DB 1; Length 396;
Best local similarity 54.5%; Pred. No. 3.9e-78;
Matches 244; Conservative 55; Mismatches 87; Indels 62; Gaps 10;
QY 3 LLARCLLVVSLVCSGLACGPRG-FGKRHPFK-LTPLAYQIPNVAEKTGASG 60
DB 7 LLPLCL-----ALLLSAQSGCGPRGPGVRRRVKQLVPLLYKQVPSMPERTLGASG 61
QY 61 RYEGKISNSRFKELPTNYPNDII FKDENTGADRLMTORCKDKNALAISVMNWPVG 120
DB 62 PAEGRVTKSGFRDLVFNYPNDII FKDENSGADRLMTCKERVNALAIIVNMWPGV 121
QY 121 KLRVTEGWDEGHSESLHYEGRVDTTSDRDSKYGMRLARLAVEAGFDWVYYSKAH 180
DB 122 RLRTVTEGWDEGHHAQDSLHYEGRALDITTSDRDNKYGLLARLAVEAGFDWVYYSRNH 181
QY 181 IHCSVKAENVAASGCGFGSATVHLEQGTGLVKDLSQDRVLAADCGRLLYSDFLT 240
DB 182 IHVSVKADNSLAVRAGCGFGNATVRLSGRKGRLRHGDWDVLAADAAGRVVFTPLL 241
QY 241 FLDRDGAQKVFYVYETREPRERLLTAHLLFVAPENDSATGFEASGSGP-PSGGAL 299
DB 242 FLDRDQRRASFVAVETERPRKULLTPEHLVFAA-----RGPAAPAGDF 286
QY 300 GPRALFASRVPGQRYVVAERDGRLLPAAVHSVTLSBEACAYAPLTAQGTILINRV 359
DB 287 AP--VFARLRAGSVLA----PGDALQPARVARVA-REBAVGVFAPLTAHGTLVNDV 339
QY 360 LASCYAVTEESWAHRAFPAPRLAHALLAALAPARTDRGDSGGGGRGGRVALTPAG 419
DB 340 LASCYAVLESQWAHRAFPAPRLHLGALL-----PG 372
QY 420 AADAPGAGATGHIWYSLLIQIGTWLL 447
DB 373 GAVQP-----TGMHWYSLRLYLAELM 395
RESULT 15
DHH_HUMAN

```

```

ID AC Q4323; Q15794; STANDARD; PRT; 396 AA.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Desert hedgehog protein precursor (DHH) (HHG-3).
GN DHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tate G., Kishimoto K., Mitsuya T.;
RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in
RL human cancer cell lines and embryonic organs.";
RN [2]
RP SEQUENCE FROM N.A.
RA Tissue=Brain;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellanc N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RN [3]
RP SEQUENCE OF 85-178 FROM N.A.
RC Tissue=Kidney;
RX Drummond I.A.;
RT "Human desert hedgehog.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP DISEASE.
RX MEDLINE=20489872; PubMed=11017805;
RA Umehara F., Tate G., Itoh K., Yamaguchi N., Douchi T., Mitsuya T.,
RA Osame M.;
RT "A novel mutation of desert hedgehog in a patient with 46,XY partial
RL gonadal dysgenesis accompanied by minifascicular neuropathy.";
CC Am. J. Hum. Genet. 67:1302-1305(2000).
CC !- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
CC DEVELOPMENT.
CC !- SUBCELLULAR LOCATION: The C-terminal peptide diffuses from the
CC cell, while the N-terminal peptide remains associated with the
CC cell surface. Is also secreted in either cleaved or uncleaved form
CC to mediate signaling to other cells (By similarity).
CC !- PTM: The C-terminal domain displays an autotransferrin activity
CC and a cholesterol transferase activity. Both activities result in
CC the cleavage of the full-length protein and covalent attachment of
CC a cholesterol moiety to the C-terminal of the newly generated N-
CC terminal fragment (N-product). This covalent modification appears
CC to play an essential role in restricting the spatial distribution
CC of the protein activity to the cell surface. The N-product is the
CC active species in both local and long-range signaling, whereas the
CC C-product has no signaling activity (By similarity).
CC !- DISEASE: Defects in DHH may be the cause of partial gonadal

```

